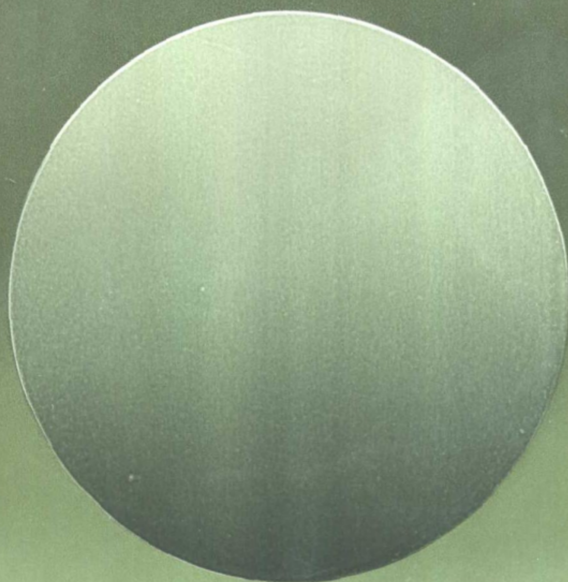
●研究生用书●

SEMIMARTINGALE'S SEQUENCES: THEORY AND ITS APPLICATIONS

华中理工大学出版社



胡必锦

半鞅序列理论及应用

ISBN 7-5609-1502-7

定价：15.00元

半鞅序列理论及应用

胡必锦

华中理工大学出版社

(鄂)新登字第10号

图书在版编目(CIP)数据
半鞅序列理论及应用 / 胡必锦 著
武汉:华中理工大学出版社,1997年3月
ISBN 7-5609-1502-7
Ⅰ. 半…
Ⅱ. 胡…
Ⅲ. 概率论-半鞅序列-研究
Ⅳ. O211

·研究生用书·
半鞅序列理论及应用
胡必锦 著
责任编辑:李立鹏
*
华中理工大学出版社出版发行
(武昌喻家山 邮编:430074)
新华书店湖北发行所经销
华中理工大学出版社照排室排版
华中理工大学出版社沔阳印刷厂印刷
*
开本:850×1168 1/32 印张:12.25 插页:2 字数:302 000
1997年3月第1版 1997年3月第1次印刷
印数:1—2 000
ISBN 7-5609-1502-7/O·166
定价:15.00元
(本书若有印装质量问题,请向出版社发行部调换)

作者简介

胡必锦，生于 1940 年 5 月。1964 年毕业于武汉大学数学系。1967 年于武汉大学数学系研究生毕业。1985～1986 年在美国德州大学雅林顿分校当访问学者，1984 年晋升为副教授。1989 年晋升为教授。目前从事统计反演方面的研究工作。

内 容 简 介

本书介绍半鞅序列的基本理论兼某些经典的应用课题，如最优停时、假设检验等．全书共分七章，前三章为基础知识，第四章介绍半鞅序列的基本概念和基础理论．第五、六两章介绍应用专题．最后为方便读者附录 Markov 链．

本书的材料选择按照循序渐进、不避重就轻的原则，自成体系．凡较重要的概念和结论均有较详尽的陈述．对于那些具有大学本科三年数学基础的读者无须额外查询资料便可顺利地阅读．该书既可作为数学专业的研究生相关内容的教学参考材料，也适宜于大学高年级学生自学．

Abstract of the text

In the course, the fundamental theory of semimartingale's sequences and its applications, such as optimal stopping time and testing hypothesis et., are introduced. The course is divided into seven chapters. The primary knowledge is described in first three chapters. Fundamental concepts and conclusions of semimartingales are set into fourth chapter of the course. Topics considered as applications of semimartingales are introduced in fiveth-sixth chapter. Finally, Markovian train is set in the appendix of the course in order to make reading the course convenient.

Materials of the course are chosen and edited in accordance with principle of order and advance step by step, and form a complete system. Important concepts and conclusions are detailed in the course. Readers with three-year mathematical foundation can read the course without any difficult. The course may be considered as a book of teaching reference and self-learned materials for the graduate students.

“研究生用书”总序

研究生教材建设是提高研究生教学质量的重要环节，是具有战略性的基本建设。各门课程必须有高质量的教材，才能使学生通过学习掌握各门学科的坚实的基础理论和系统的专门知识，为从事科学研究工作或独立担负专门技术工作打下良好的基础。

我校各专业自 1978 年招收研究生以来，组织了一批学术水平较高，教学经验丰富的教师，先后编写了公共课、学位课所需的多种教材和教学用书。有的教材和教学用书已正式出版发行，更多则采用讲义的形式逐年印发。这些讲义经过任课教师多年教学实践，不断修改、补充、完善，已达到出书的要求。因此，我校决定出版“研究生用书”，以满足本校各专业研究生教学需要，并与校外单位交流，征求有关专家学者和读者的意见，以促进我校研究生教材建设工作，提高教学质量。

“研究生用书”以公共课和若干门学位课教材为主，还有教学参考书和学术专著，涉及的面较广，数量较多，准备在今后数年内分批出版。编写“研究生用书”总的要求是从研究生的教学需要出发，根据各门课程在教学过程中的地位和作用，在内容上求新、求深、求精，每本教材均应包括本门课程的基本内容，使学生能掌握必需的基础理论和专门知识；学位课教材还应接触该学科的发展前沿，反映国内外的最新研究成果，以适应目前科学技术知识更新很快的形势；学术专著则应充分反映作者的科

研硕果和学术水平，阐述自己的学术见解。在结构和阐述方法上，应条理清楚，论证严谨，文字简炼，符合人们的认识规律。总之，要力求使“研究生用书”具有科学性、系统性和先进性。

我们的主观愿望虽然希望“研究生用书”的质量尽可能高一些，但由于研究生的培养工作为时尚短，水平和经验都不够，其中缺点、错误在所难免，尚望校内外专家学者及读者不吝指教，我们将非常感谢。

华中理工大学研究生院院长

陳 珽 黄樹槐

1989.11.

写在 1995 年

在今天，国家之间的竞争是国家综合实力的竞争，国家综合实力的竞争关键是经济实力的竞争，而经济实力的竞争关键又在于科技（特别是高科技）的竞争，科技（特别是高科技）的竞争归根结底是人才（特别是高层次人才）的竞争，而人才（特别是高层次人才）的竞争基础又在于教育。"百年大计，教育为本；国家兴亡，人才为基。"十六个字、四句话，确是极其深刻的论断。

显然，作为高层次人才培养的研究生教育就在一个国家的方方面面的工作中，占有十分重要的战略地位。可以说，没有研究生教育，就没有伟威雄壮的科技局面，就没有国家的强大实力，就没有国家在国际上的位置，就会挨打，就会受压，就会被淘汰。

"工欲善其事，必先利其器。"教学用书是教学的重要基本工具与条件。这是所有从事教育的专家所熟知的事实。所以，正如许多专家所知，也正是原来的《"研究生用书"总序》中所指出，研究生教材建设是保证与提高研究生教学质量的重要环节，是一项具有战略性的基本建设。没有研究生的质量，就没有研究生教育的一切。

我校从 1978 年招收研究生以来，即着力从事于研究生教材与教学用书的建设。积十多年建设与实践的经验，我校从 1989 年起，正式分批出版"研究生用书"。第一任

研究生院院长陈珽教授就为之写了《"研究生用书"总序》,表达了我校编写这套用书的指导思想与具体要求,"要力求'研究生用书'具备科学性、系统性、先进性"。第二任研究生院长,也就是当时我校的校长黄树槐教授完全赞同这一指导思想与具体要求,从多方面对这套用书加以关心与支持。

我是十分支持出版"研究生用书"的。早在1988年我在为列入这套书中的第一本,即《机械工程测试·信息·信号分析》写"代序"时就提出:一个研究生应该博览群书,博采百家,思路开阔,有所创见。但这不等于他在一切方面均能如此,有所不为才能有所为。如果一个研究生的主要兴趣与工作不在"这一特定方面",他也可以选择一本有关的书作为了解与学习这方面专业知识的参考;如果一个研究生的主要兴趣在"这一特定方面",他更应选择一本有关的书作为主要学习用书,寻觅主要学习线索,并缘此展开,博览群书。这就是我赞成为研究生编写系列教学用书的原因。

目前,这套用书已出版了6批共30种,初步形成规模,逐渐为更多读者所认可。在已出版的书中,有8种分获国家级、部省级图书奖,有13种一再重印,久销不衰,有的印刷总数已近万册。采用此套书的一些兄弟院校教师纷纷来信,赞誉此书为研究生培养与学科建设作出了贡献,解决了他们的"燃眉之急"。我们感谢这些赞誉与鼓励,并将这些作为对我们的鞭策与鼓励,"衷心藏之,何日忘之?!"

现在,正是江南初春,"最是一年春好处"。华工园内,

红梅怒放，迎春盛开，柳枝抽绿，梧叶含苞，松柏青翠，樟桂换新，如同我们的国家正在迅猛发展，欣欣向荣一样，一派盎然生机。尽管春天还有乍寒时候，我们国家在前进中还有种种困难与险阻，有的还很严峻，但是，潮流是不可阻挡的，春意会越来越浓，国家发展会越来越好。我们教师所编的、所著的、所编著的这套教学用书，也会在解决前进中的种种问题中继续发展。然而，我们十分明白，这套书尽管饱含了我们教师的辛勤的长期的教学与科研工作的劳动结晶，作为教学用书百花园中的一丛鲜花正在怒放，然而总会有这种或那种的不妥、错误与不足，我衷心希望在这美好的春日，广大的专家与读者，不吝拨冗相助，对这套教学用书提出批评建议，予以指教启迪，为这丛鲜花除害灭病，抗风防寒，以进一步提高质量，提高水平，更上一层楼，我们不胜感激。我们深知，"一个篱笆三个桩"，没有专家的指导与支持，没有读者的关心与帮助，也就没有这套教学用书的今天。

诗云："嘤其鸣矣，求其友声。"这是我们的心声。

中国科学院院士
华中理工大学校长
兼研究生院院长

杨叔子
于华工园内
1995年3月7日

前　言

随机序列在概率论中是一类特别重要的特殊随机过程，比如：随机游动、Markov 链、半鞅序列等. 本书将以半鞅序列及停时理论为主要讨论对象，作为一本普及性的参考书，考虑到在校生及自学者等不同层面的需要，我们在选材上侧重介绍系统的理论基础及某些实用性专题. 如果读者具备本科大学三年级的基础知识，则不必依靠其它参考资料便可顺利地阅读.

本书分七章，前四章及最后的附录为本书的基础部分. 第五章介绍最优停时理论(周金海、刘丁酉参与这章的编写工作). 第六章介绍假设检验的基本知识，侧重讨论鞅与停时理论在这个专题研究中的应用. 每一章除它们之间的联系外还具有一定的独立性和完整性. 最后列出了一些参考文献.

本书在编写过程中得到华中理工大学研究生院的资助以及华中理工大学出版社的支持，武大数学系胡迪鹤教授对本书提出了许多宝贵的修改意见. 在此，一并表示感谢. 本书内容涉及面较广，限于作者的水平，错漏难免，敬请专家、读者批评指正.

作者

1996. 6.

目　　录

第一章 一般测度

§1.1 集 类

设 Ω 为某个基本事件空间. $\varnothing$表示空集.

定义 1.1.1(半环) 设 $\mathscr{A}$ 为 Ω 上的一个子集类. 若

(i) $\varnothing\in\mathscr{A}$; (ii) $A,B\in\mathscr{A}\Rightarrow A\cap B\in\mathscr{A}$;

(iii) $A,B\in\mathscr{A}$, 且 $B\subset A\Rightarrow\exists\ C_i\in\mathscr{A}$, $C_i\cap C_j=\varnothing(i\neq j)$及某整数 $n\geqslant1$, 使 $A\setminus B=\bigcup\limits_{i=1}^{n}C_i$,

则称 $\mathscr{A}$ 为 Ω 上的半环.

引理 1.1.1 设 $\mathscr{A}$ 为 Ω 上的半环, $E_i\in\mathscr{A}(i\geqslant1)$, 则 $\bigcup\limits_i E_i=\bigcup\limits_k C_k$, 其中 $C_k\in\mathscr{A}$, $C_i\cap C_k=\varnothing(i\neq k)$.

证 令 $\widetilde{E}_1=E_1$, $\widetilde{E}_m=E_m\setminus\bigcup\limits_{i=1}^{m-1}\widetilde{E}_i(m\geqslant2)$, 则

$$\bigcup_i E_i=\bigcup_i\widetilde{E}_i,\quad \widetilde{E}_i\cap\widetilde{E}_j=\varnothing(i\neq j).$$

因此, 命题等价于: 对$\forall\ E\in\mathscr{A}$, $m\geqslant1$, 有

$$E\setminus\bigcup_{i=1}^{m}E_i=\bigcup_{l=1}^{m^*}C_l, \tag{1.1-1}$$

其中 $C_l\in\mathscr{A}$, $C_l\cap C_k=\varnothing(l\neq k)$, $m^*\geqslant1$.

当 $m=1$ 时, $E\setminus E_1=E\setminus E\cap E_1$. 按半环条件(ii)和(iii), (1.1-1)式成立. 现假定(1.1-1)式对 $m\geqslant1$ 真, 则

$$E\setminus\bigcup_{i=1}^{m+1}E_i=E\setminus E_{m+1}\setminus\bigcup_{i=1}^{m}E_i=\Big(\bigcup_{k=1}^{n}\widetilde{C}_k\Big)\setminus\bigcup_{i=1}^{m}E_i=\bigcup_{k=1}^{n}\Big(\widetilde{C}_k\setminus\bigcup_{i=1}^{m}E_i\Big).$$

按假定, 对给定的 $m\geqslant1$, 有

$$\widetilde{C}_k\setminus\bigcup_{i=1}^{m}E_i=\bigcup_{l=1}^{m_k}C_l^{(k)}=\bigcup_{l=1}^{m^*}C_l^{(k)}.$$

其中 $$m^* = \max_{1\leqslant k\leqslant n}\{m_k\},\ C_l^{(k)} = \varnothing\ (l > m_k).$$

令 $C_l = C_r^{(k)}$，其中 $l = (k-1)m^* + r\ (1\leqslant k\leqslant n,\ 1\leqslant r < m^*)$，则

$$E\setminus\bigcup_{i=1}^{m+1} E_i = \bigcup_{l=1}^{(m+1)^*} C_l,\quad \text{其中}(m+1)^* = n\cdot m^*.$$

定义 1.1.2（环和代数） 设 $\mathscr{A}$ 为 Ω 上的一个子集类. 若

(i) $\mathscr{A}$ 非空，

(ii) $A, B\in\mathscr{A}\Rightarrow A\cup B\in\mathscr{A},\ A\setminus B\in\mathscr{A}$.

则称 $\mathscr{A}$ 为 Ω 上的**环**.

若 $\mathscr{A}$ 为 Ω 上的环，且 $\Omega\in\mathscr{A}$，则称 $\mathscr{A}$ 为 Ω 上的**代数**.

注 1 有时称含 Ω 的半环为**半代数**.

注 2 环 $\mathscr{A}$ 中的条件(ii)等价于如下条件：

(ii)′ $A, B\in\mathscr{A}\Rightarrow A\cap B\in\mathscr{A},\ A\triangle B\in\mathscr{A}$（"$\triangle$"表示对称差）.

定义 1.1.3（单调类） 设 $\mathscr{A}$ 为 Ω 上的非空子集类. 若 $(A_k, k\geqslant 1)$ 为 $\mathscr{A}$ 中的单调子列，且 $A = \lim\limits_{n\to\infty} A_n\in\mathscr{A}$，则称 $\mathscr{A}$ 为 Ω 上的**单调类**.

定义 1.1.4（π-类，λ-类） 设 $\mathscr{A}$ 为 Ω 上的非空子集类. 若 $A, B\in\mathscr{A}\Rightarrow A\cap B\in\mathscr{A}$，则称 $\mathscr{A}$ 为 Ω 上的 **π-类**. 若(i) $\Omega\in\mathscr{A}$；(ii) $A\in\mathscr{A}\Rightarrow A^c\in\mathscr{A}$；(iii) $(A_n, n\geqslant 1)\in\mathscr{A}$，且 $A_n\cap A_m = \varnothing\ (n\neq m)\Rightarrow\bigcup_n A_n\in\mathscr{A}$. 则称 $\mathscr{A}$ 为 Ω 上的 **λ-类**.

定义 1.1.5 设 $\mathscr{A}_0$ 为 Ω 上的非空子集类. β 表示某种特性. 记

$$\beta(\mathscr{A}_0) = \inf\{\mathscr{A}:\ \mathscr{A}_0\subset\mathscr{A},\ \mathscr{A}\text{ 具有性质 }\beta\},$$

则称 $\beta(\mathscr{A}_0)$ 为 Ω 上含 $\mathscr{A}_0$，且具有性质 β 的**最小集类**.

例如，$\pi(\mathscr{A}_0)$，$\lambda(\mathscr{A}_0)$，$m(\mathscr{A}_0)$，$r(\mathscr{A}_0)$，$\alpha(\mathscr{A}_0)$，$\sigma(\mathscr{A}_0)$分别表示含 $\mathscr{A}_0$ 的最小 π-类，λ-类，单调类，环，代数，σ-代数.

引理 1.1.2 设 $\mathscr{A}$ 为 Ω 上的子集类.

(i) 若 $\mathscr{A}$ 为代数，则 $\mathscr{A}$ 为 σ-代数的充要条件是

$$\mathscr{A} = m(\mathscr{A}).$$

(ii) 若 $\mathscr{A}=\pi(\mathscr{A})$，则 $\lambda(\mathscr{A})=\sigma(\mathscr{A})$.

证 结论(i)之证留给读者. 下证结论(ii).

假如 $\lambda(\mathscr{A})$ 为 σ-代数，则 $\sigma(\mathscr{A})\subset\lambda(\mathscr{A})$，但 $\sigma(\mathscr{A})$ 为含 $\mathscr{A}$ 的 λ-类. 因此，$\lambda(\mathscr{A})\subset\sigma(\mathscr{A})$，故 $\lambda(\mathscr{A})\subset\sigma(\mathscr{A})$. 剩下的问题是证明：$\lambda(\mathscr{A})$ 为 σ-代数. 令

$$\mathscr{G}_A=\{B:A\cap B\in\lambda(\mathscr{A})\},$$

则对 $\forall A\in\lambda(\mathscr{A})$，$\mathscr{G}_A$ 为 λ-类. 事实上，由 $A\cap\Omega=A$ 可知：$\Omega\in\mathscr{G}_A$. 若 $B\in\mathscr{G}_A$，则 $A\cap B^c=(A^c\cup A\cap B)^c\in\lambda(\mathscr{A})$. 从而，$B^c\in\mathscr{G}_A$. 假如，$(B_n,\ n\geqslant1)\subset\mathscr{G}_A$，且 $B_n\uparrow B$，则

$$(A\cap B_n,\ n\geqslant1)\subset\lambda(\mathscr{A}),\ 且\ A\cap B_n\uparrow A\cap B.$$

于是，$A\cap B\in\lambda(\mathscr{A})\Rightarrow B=\lim\limits_{n\to\infty}B_n\in\mathscr{G}_A$. 按定义 1.1.5，$\mathscr{G}_A$ 为 λ-类.

显然，$\mathscr{A}\subset\mathscr{G}_A(\forall A\in\mathscr{A})$，$\lambda(\mathscr{A})\subset\mathscr{G}_A(\forall A\in\lambda(\mathscr{A}))$. 从而，$\mathscr{A}\subset\mathscr{G}_A$，$\lambda(\mathscr{A})\subset\mathscr{G}_A(\forall A\in\lambda(\mathscr{A}))$. 于是，若 $A,B\in\lambda(\mathscr{A})$，则 $A\in\mathscr{G}_B\Rightarrow A\cap B\in\lambda(\mathscr{A})$，即 $\lambda(\mathscr{A})$ 为 π-类. 易证：既为 λ-类又为 π-类的集类必为 σ-代数，故 $\lambda(\mathscr{A})$ 为 σ-代数.

定理 1.1.1(Dykinπ-λ 定理) 设 $\mathscr{P}$ 为 π-类，$\mathscr{L}$ 为 λ-类，且 $\mathscr{P}\subset\mathscr{L}$，则 $\sigma(\mathscr{P})\subset\mathscr{L}$.

证 按引理 1.1.2(ii)，$\sigma(\mathscr{P})=\lambda(\mathscr{P})\subset\mathscr{L}$.

在定义 1.1.5 中，$\sigma(\mathscr{A})$ 表示含 $\mathscr{A}$ 的最小 σ-代数. 假如 $\mathscr{A}$ 是由有限个元素构成的集类，则易证：$\sigma(\mathscr{A})$ 可以通过有限次集论运算(补、可数并(交)等)构造出来. 然而，对一般的子集类 $\mathscr{A}$，则不一定能通过有限次集论运算构作出 $\sigma(\mathscr{A})$. 下面讨论一个反例.

设 $\mathscr{H}$ 为 Ω 上的某个集类. 记

$$\mathscr{H}^*=\{A:A\ 由\ \mathscr{H}\ 中的元素通过有限次集论运算产生的结果\}.$$

令 $\mathscr{A}_1=\mathscr{H}$，$\mathscr{A}_k=\mathscr{A}_{k-1}^*(k\geqslant2)$，则 $\mathscr{A}_n\subset\sigma(\mathscr{H})(\forall n\geqslant1)$. 但一般没有 $\bigcup\limits_n\mathscr{A}_n=\sigma(\mathscr{H})$. 下列命题将证明这一点.

命题 1.1.1 设 $\Omega=(0,1]$，

$$\mathscr{I}_0 = \{(a,b]: a,b \text{ 为}[0,1] \text{ 中的有理数,且 } a \leqslant b\},$$

$$\mathscr{I}_n = \mathscr{I}_{n-1}^{*} (\forall\, n \geqslant 1).$$

则$\bigcup\limits_n \mathscr{I}_n$严格地被含在$\sigma(\mathscr{I}_0)$中.

证 首先讨论$\bigcup\limits_n \mathscr{I}_n$中元素的表示. 令

$$\Psi(A_1, A_2, \cdots) = A_1^c \cup A_2 \cup A_3 \cup \cdots$$

则对$\forall\, B^n \in \mathscr{I}_n$,有$B^n = \Psi(A_1, A_2, \cdots)$,其中$A_k \in \mathscr{I}_{n-1} (k \geqslant 1)$. 这一事实可直接由$\mathscr{I}_n = \mathscr{I}_{n-1}^{*}$得到. 假定每个正整数在正整数方阵$(m_{ij})_{i,j\geqslant 1}$中恰好只出现一次. 定义

$$\Phi_0(A_1, A_2, \cdots) = A_1; \text{ 对 } n \geqslant 1,$$

$$\Phi_n(A_1, A_2, \cdots) = \Psi(\Phi_{n-1}(A_{m_{11}}, A_{m_{12}}, \cdots),$$
$$\Phi_{n-1}(A_{m_{21}}, A_{m_{22}}, \cdots), \cdots),$$

则采用归纳法可以证明:对$\forall\, B^n \in \mathscr{I}_n$,都有表示

$$B^n = \Phi_n(A_1, A_2, \cdots), \quad (A_k,\ k \geqslant 1) \subset \mathscr{I}_0.$$

令 $\Phi(A_1, A_2, \cdots) = \bigcup\limits_n \Phi_n(A_1, A_2, \cdots) \quad ((A_k,\ k \geqslant 1) \subset \mathscr{I}_0),$

则对$\forall\, B \in \bigcup\limits_n \mathscr{I}_n$,有表示

$$B = \Phi(A_1, A_2, \cdots) \quad ((A_k,\ k \geqslant 1) \subset \mathscr{I}_0).$$

按归纳法,立即可证:$\Phi(A_1, A_2, \cdots) \in \sigma(\mathscr{I}_0)$, $((A_k,\ k \geqslant 1) \subset \mathscr{I}_0)$.
$\Rightarrow \bigcup\limits_n \mathscr{I}_n \subset \sigma(\mathscr{I}_0)$.

现在,来构作一个 Borel 集B不含在$\bigcup\limits_n \mathscr{I}_n$中. 为此,让$\omega \in \Omega$,则相应于$\omega$有一个正整数列$(\omega_1, \omega_2, \cdots)$,使得$\omega$的二进制展开式中第$k$个"1"的位置编号为$\omega_1 + \omega_2 + \cdots + \omega_k$,即令$n = \sum\limits_{i=1}^{k} \omega_i$,则

$$n = \inf\left\{m:\ \sum_{i=1}^{k} \omega_i = m,\ \sum_{j=1}^{m} d_j(\omega) = k\right\},$$

其中$d_j(\omega) = 0$或1为ω的二进制展式中第j个位置的数字. 在此意义下,$(\omega_k,\ k \geqslant 1)$唯一被决定. 于是,$\omega \in (0,1]$与$(\omega_k,\ k \geqslant 1)$是1-1对应的.

设$(I_k, k\geqslant 1)\subset\mathscr{I}_0$, $\varphi(\omega)=\Phi(I_{\omega_1}, I_{\omega_2}, \cdots)$, $B=\{\omega: \omega\notin\varphi(\omega)\}$. 则 B 为 Borel 集，且不含在任何 $\mathscr{I}_n$ 中. 事实上，对$\forall\ \omega\in(0,1]$，有 $B\neq\varphi(\omega)$. 而 $\varphi(\omega)\in\bigcup_n\mathscr{I}_n$，且$\bigcup_n\mathscr{I}_n$ 中的每个元素对应一个 $\omega\in(0,1]$和$(I_k, k\geqslant 1)\subset\mathscr{I}_0$，故 $B\notin\bigcup_n\mathscr{I}_n$. 现在，令

$$D_k=\{\omega: \omega\in I_{\omega_k}\},$$

$$L_k(n)=\left\{\omega: \sum_{i=1}^{k}\omega_i=n\right\}=\left\{\omega: \sum_{j=1}^{n-1}d_j(\omega)<k=\sum_{j=1}^{n}d_j(\omega)\right\}.$$

则显然有 $L_k(\omega)\in\sigma(\mathscr{I}_0)$.

$$\{\omega: \omega_k=n\}=\bigcup_{m=1}^{\infty}L_{k-1}(m)\cap L_k(m+n)\in\sigma(\mathscr{I}_0),$$

$$D_k=\bigcup_{n=1}^{\infty}\{\omega: \omega_k=n\}\cap I_n\in\sigma(\mathscr{I}_0).$$

假如，$\{\omega: \omega\in\Phi_n(I_{\omega_{u_1}}, I_{\omega_{u_2}}, \cdots)\}=\Phi_n(D_{u_1}, D_{u_2}, \cdots)$, (1.1-2)

其中 $n\in\mathrm{N}$，$(u_k, k\geqslant 1)$为任意正整数列，则

$$B^c=\{\omega: \omega\in\varphi(\omega)\}=\bigcup_{n=1}^{\infty}\{\omega: \omega\in\Phi_n(I_{\omega_{m_{n_1}}}, I_{\omega_{m_{n_2}}}, \cdots)\}$$

$$=\bigcup_{n=1}^{\infty}\Phi_n(D_{m_{n_1}}, D_{m_{n_2}}, \cdots)=\Phi(D_1,D_2,\cdots)\in\sigma(\mathscr{I}_0).$$

$\Rightarrow B\in\sigma(\mathscr{I}_0)$，即 B 为 Borel 集.

剩下的问题就是证明(1.1-2)式成立.

采用归纳法，当 $n=0$ 时，

$$\{\omega: \omega\in\Phi_0(I_{\omega_{u_1}}, I_{\omega_{u_2}}, \cdots)\}=\{\omega: \omega\in I_{\omega_{u_1}}\}$$

$$=D_{u1}=\Phi_0(D_{u_1}, D_{u_2}, \cdots).$$

这表明 $n=0$ 时，(1.1-2)式成立. 假如(1.1-2)式对 $n-1$ 成立. 并设 $\omega\in\Phi_n(I_{\omega_{u_1}}, I_{\omega_{u_2}}, \cdots)$，则由$(\Phi_n, n\geqslant 0)$和 Ψ 之定义可知：存在某 $k\geqslant 1$，使得

$$\omega\in\Phi_{n-1}(I_{\omega_{u_{m_{k_1}}}}, I_{\omega_{u_{m_{k_2}}}}, \cdots).$$

按归纳法假设，得：

$$\Phi_n(I_{\omega_{u_1}}, I_{\omega_{u_2}}, \cdots) = \Psi(\Phi_{n-1}(D_{u_{m_{11}}}, D_{u_{m_{12}}}, \cdots),$$
$$\Phi_{n-1}(D_{u_{m_{21}}}, D_{u_{m_{22}}}, \cdots), \cdots)$$
$$= \Phi_n(D_{u_1}, D_{u_2}, \cdots).$$

$\Rightarrow$(1.1-2)式对 n 亦成立.

§1.2 可测空间

设 $\mathscr{F}$ 为 Ω 上的一个 σ-代数，二重符号$(\Omega,\mathscr{F})$称为可测空间. 本节将介绍几种常用可测空间.

(一) 可测空间$(\mathrm{R},\mathscr{B})$

设 $\mathrm{R}=(-\infty,\infty)$. $[a,b]=\{x\in\mathrm{R}: a<x\leqslant b\}$, $(a,\infty]=(a,\infty)$.

$$\mathscr{A}_0=\{A: A=(a,b];\ a,b\in\overline{\mathrm{R}}=[-\infty,\infty],\ a\leqslant b\}.$$

显然，$\mathscr{A}_0$ 为半代数，且

$$\alpha(\mathscr{A}_0)=\{A: A=\bigcup_{i=1}^{n}A_i;\ 1\leqslant n<\infty,$$
$$A_i\in\mathscr{A}_0,\ A_i\cap A_j=\varnothing(i\neq j)\}.$$

这里 $\alpha(\mathscr{A}_0)$表示含 $\mathscr{A}_0$ 的最小代数. 令 $\mathscr{B}=\sigma(\mathscr{A}_0)$，则

$$\mathscr{B}=m(\alpha(\mathscr{A}_0)),$$

其中 $m(\cdot)$表示含"$\cdot$"的最小单调类. 通常称此集类为 R 上的 Borel σ-代数. 而$(\mathrm{R},\mathscr{B})$为一维 Borel 可测空间. Borel σ-代数还可以通过在 R 上引入某个拓扑结构产生. 例如：在 R 上定义尺度

$$\rho(x,y)=\frac{|x-y|}{1+|x-y|}.$$

记 $\mathscr{E}_0=\{A: A=S_a(x);\ x\in\mathrm{R},\ 0<a<1\}$,

$$S_a(x)=\{y\in\mathrm{R}: \rho(y,x)<a\}.$$

则可证明：$\mathscr{B}=\sigma(\mathscr{E}_0)$.

（二）可测空间$(\mathbf{R}^n, \mathscr{B}^n)$（$n$ 为正整数）

$$设\mathbf{R}^n=\overset{n}{\underset{1}{\times}}\mathbf{R}=\{x: x=(x_1,x_2,\cdots,x_n),\ 其中\ x_i\in\mathbf{R}\},$$

$$\mathscr{A}_0^n=\overset{n}{\underset{1}{\times}}\mathscr{A}_0=\{B^n: B^n=B_1\times\cdots\times B_n, 其中\ B_k\in\mathscr{A}_0\},$$

$$\alpha(\mathscr{A}_0^n)=\{B^n: B^n=\bigcup_{k=1}^{m}E_k;\ 其中\ m\geqslant 1,$$

$$E_k\in\mathscr{A}_0^n,\ E_k\cap E_l=\varnothing(k\neq l)\}.$$

$$\mathscr{B}^n=m(\alpha(\mathscr{A}_0^n))=\sigma(\mathscr{A}_0^n).$$

通常称$(\mathbf{R}^n, \mathscr{B}^n)$为$n$-维 Borel 可测空间. 从构作中立即可以看出：此可测空间为一维 Borel 可测空间的乘积空间. 有时记为 $\mathscr{B}^n=\overset{n}{\underset{1}{\times}}\mathscr{B}$；$(\mathbf{R}^n, \mathscr{B}^n)=\overset{n}{\underset{1}{\times}}(\mathbf{R}, \mathscr{B})$.

类似于一维情形，$\mathscr{B}^n$ 也可以通过在 $\mathbf{R}^n$ 中引入某个拓扑结构产生，如在 $\mathbf{R}^n$ 上引入尺度：

$$\rho_n(x,y)=\sum_{k=1}^{n}2^{-n}\rho(x_k,y_k)\quad(x,y\in\mathbf{R}^n, x=(x_1,x_2,\cdots,x_n)).$$

相应此尺度有如下邻域系：

$$\mathscr{E}_0^n=\{B: B=S_a^n(x),\ 其中\ x\in\mathbf{R}^n,\ 0<a<1\};$$

$$S_a^n(x)=\{y\in\mathbf{R}^n: \rho_n(x,y)<a\}.$$

同样可证：$\mathscr{B}^n=\sigma(\mathscr{E}_0^n)$.

（三）可测空间$(\mathbf{R}^\infty, \mathscr{B}^\infty)$

$$设\mathbf{R}^\infty=\overset{\infty}{\underset{1}{\times}}\mathbf{R}=\{x: x=(x_1,x_2,\cdots),\ x_k\in\mathbf{R}\},$$

$$\mathscr{B}^\infty=\overset{\infty}{\underset{1}{\times}}\mathscr{B}=\sigma(\{B^\infty: B^\infty=\overset{\infty}{\underset{i=1}{\times}}B_i,\ B_i\in\mathscr{B}\}),$$

$$\mathscr{A}_0^\infty=\{A: A=\mathscr{I}(I_1\times I_2\times\cdots\times I_n),\ 其中\ n\geqslant 1,\ I_k\in\mathscr{A}_0\},$$

$$\mathscr{I}(\overset{\infty}{\underset{k=1}{\times}}I_k)=\{x\in\mathbf{R}^\infty: x_k\in I_k,\ 1\leqslant k\leqslant n\}.$$

则 $\mathscr{B}^\infty=\sigma(\mathscr{A}_0^\infty)$. 通常称 $\mathscr{B}^\infty$ 为无穷维 Borel σ-代数. $(\mathbf{R}^\infty, \mathscr{B}^\infty)$ 为无穷维 Borel 可测空间.

σ-代数 $\mathscr{B}^\infty$ 也可以通过引入某种拓扑结构产生. 如在 R^∞ 上引入尺度

$$\rho_\infty(x,y) = \sum_{k=1}^{\infty} 2^{-k}\rho(x_k, y_k) \quad (x, y \in \mathrm{R}^\infty).$$

令
$$D_a(x) = \{y \in \mathrm{R}^\infty:\ \rho_\infty(y,x) < a\},$$
$$\mathscr{E}_0^\infty = \{D_a(x):\ x \in \mathrm{R}^\infty,\ 0 < a < 1\}.$$

则
$$\mathscr{B}^\infty = \sigma(\mathscr{E}_0^\infty).$$

（四）可测空间$(\mathrm{R}^T, \mathscr{B}^T)$$(T=[a,b]$或$[0,\infty))$

设
$$\mathrm{R}^T=\{x:\ x=(x_t,\ t\in T),\ x_t\in \mathrm{R}\},$$
$$\mathscr{I}^T_{t_1\cdots t_n}(B^n)=\{x\in \mathrm{R}^T:\ (x_{t1},\ x_{t2},\ \cdots,\ x_{tn})\in B^n\},$$
$$\mathscr{A}^T_0=\{A:\ A=\mathscr{I}^T_{t_1,\cdots,t_n}\Big(\mathop{\times}_{k=1}^{n} I_k\Big),\ n\geqslant 1,\ t_k\in T,\ t_1<\cdots<t_n,$$
$$I_k\in \mathscr{A}_0\},$$
$$\mathscr{B}^T=\sigma(\mathscr{A}^T_0)=m(\alpha(\mathscr{A}^T_0)).$$

一般称$(\mathrm{R}^T, \mathscr{B}^T)$为具有参数集 T 的乘积空间. $\mathscr{B}^T$ 中的元素有如下表现定理.

定理 1.2.1 设 T 为一个不可数的有序集,则对$\forall\ A\in \mathscr{B}^T$, $\exists\ T$ 中一个可数集$(t_k,\ k\geqslant 1)$及元素 $B\in \mathscr{B}^\infty$, 使

$$A = \{x \in \mathrm{R}^T:\ (x_{t_k},\ k \geqslant 1) \in B\}.$$

证 设 $\mathscr{E}=\{A:\ A=\mathscr{I}^T_{t_1t_2\cdots}(B),\ (t_k,\ k\geqslant 1)\subset T,\ B\in \mathscr{B}^\infty\}$. 若

$$\mathscr{E} = \mathscr{B}^T. \tag{1.2-1}$$

则定理成立. 下面证明(1.2-1)式真.

为此,首先证明 $\mathscr{E}$ 为 σ-代数. 假定$(A_k,\ k\geqslant 1)\subset \mathscr{E}$,则 $\mathscr{E}$ 之定义, $A_k=\mathscr{I}^T_{T_k}(B_k)$,其中

$$T_k = (t_i^k,\ i \geqslant 1), \quad B_k \in \mathscr{B}^\infty.$$

令 $T^* = \bigcup_{k=1}^{\infty} T_k$, $B = \bigcup_{k=1}^{\infty} \hat{B}_k$, $\hat{B}_k = \mathscr{I}^{T^*}_{T_k}(B_k) \in \mathscr{B}^\infty$, 则 $A_k = \mathscr{I}^T_{T^*}(\hat{B}_k)$. 从而,

$$\bigcup_{k=1}^{\infty} A_k = \bigcup_{k=1}^{\infty} \mathscr{I}_{T^*}^{T}(\hat{B}_k) = \mathscr{I}_{T^*}^{T}(\bigcup_{k=1}^{\infty} \hat{B}_k) = \mathscr{I}_{T^*}^{T}(B).$$

注意,T^* 可数,且 $B\in\mathscr{B}^{\infty}$,故$\bigcup_{k=1}^{\infty} A_k\in\mathscr{E}$.

$\Rightarrow\mathscr{E}$ 为 σ-代数.

显然,$\mathscr{A}_0^T\subset\mathscr{E}$,于是,$\mathscr{B}^T=\sigma(\mathscr{A}_0^T)\subset\mathscr{E}$. 按 $\mathscr{E}$ 之定义,有 $\mathscr{E}\subset\mathscr{B}^T$. $\Rightarrow$ (1.2.1)式真.

注 1 若 T 为可数集或有限集,则上述定理自然成立. 因此,定理中的 T 可以为任意有序集.

注 2 σ-代数 $\mathscr{B}^T$ 是由柱集类产生的,且按上述定理,它的每个元素都是具有以无穷维 Borel 集为基的柱集. 这个事实表明:R^T 中不属于 $\mathscr{B}^T$ 的子集就一定没有定理中所述的表现. 例如,设 $T=[0,1]$,$c\in\mathrm{R}$. 考虑子集 $A^*=\{x\in\mathrm{R}^T: \sup\{x_t: t\in T\}<c\}$,则 $A^*\notin\mathscr{B}^T$. 事实上,假若 $A^*\in\mathscr{B}^T$,则按上述定理,有

$$\exists\ (t_k,\ k\geqslant 1)\subset T,\ B\in\mathscr{B}^{\infty},$$

$$\ni A^* = \{x\in\mathrm{R}^T:\ (x_{t_k},\ k\geqslant 1)\in B\}.$$

令 $y_t=c-1(\forall\ t\in T)$,则 $y\in A^*$. 定义

$$z_t=\begin{cases} c+1, & t\notin(t_k,\ k\geqslant 1), \\ c-1, & t\in(t_k,\ k\geqslant 1). \end{cases}$$

则 $z=(z_t,\ t\in T)\notin A^*$,但 $z_t=y_t(\forall\ t\in(t_k,\ k\geqslant 1))$. 这表明:$z\in A^*$. 于是,得到矛盾的结果. 然而,像 A^* 这样的集在过程论中是相当重要的. 为了保证这类集的可测性,往往需要考虑某些比 R^T 更小的函数空间. 比如连续函数空间,右连左极函数空间等.

(五) 可测空间$(C,\ \mathscr{B}(C))$

设 $T=[0,1]$,$C=\{x\in\mathrm{R}^T:\ x=(x_t,\ t\in T)$ 为 T 上的连续函数$\}$. 在 C 上赋予如下尺度:

$$\rho_0(x,y)=\sup\{|x_t-y_t|:\ t\in T\}.$$

假定$\mathscr{B}(C)$是由 C 中的柱集类产生的 σ-代数,$\mathscr{E}_0$ 为 C 上关于尺度 ρ_0 产生的邻域系,则

$$\mathscr{B}(C) = \sigma(\mathscr{E}_0).$$

事实上，令

$$B_{t_0}^{(b)} = \{x \in C: x_{t_0} < b\} \quad (t_0 \in T,\ b \in \mathrm{R}).$$

则 $B_{t_0}^{(b)}$ 为 C 中的一个柱集，且

$$\mathscr{B}(C) = \sigma(B_t^{(b)}: t \in T,\ b \in \mathrm{R}).$$

注意，对$\forall\ x \in B_t^{(b)}$，$\exists\ 0<\delta<\frac{1}{2}(b-x_t)$， 于

$$G_\delta(x) = \{y \in C: \rho_0(y,x) < \delta\} \subset B_t^{(b)}.$$

$$\Rightarrow B_t^{(b)} \in \sigma(\mathscr{E}_0) \Rightarrow \mathscr{B}(C) \subset \sigma(\mathscr{E}_0).$$

反之，由连续性，有

$$\begin{aligned}\overline{G}_\delta(x) &= \{y \in C: \sup\{|y_t - x_t|: t \in Q\} \leqslant \delta\} \\ &= \bigcap_{t\in Q}\{y \in C: |y_t - x_t| \leqslant \delta\} \in \mathscr{B}(C),\end{aligned}$$

其中 Q 为 T 中有理数全体，不难证明：

$$\sigma(\mathscr{E}_0) = \sigma(\{\overline{G}_\delta(x): x \in C,\ 0 < \delta < 1\}).$$

$$\Rightarrow \sigma(\mathscr{E}_0) \subset \mathscr{B}(C).$$

如果在注 2 中，将 A^* 中的 R^T 换成 C，则 $A^* \in \mathscr{B}(C)$，即 A^* 为可测空间$(C,\ \mathscr{B}(C))$上的可测集.

（六）可测空间$(D,\ \mathscr{B}(D))$

设 $T=[0,1]$，$x_{t+}=\lim\limits_{s\downarrow t}x_s$，$x_{t-}=\lim\limits_{s\uparrow t}x_s$ $(0<t\leqslant 1)$. $x_{0-}=x_{0+}$.

$$D = \{x \in \mathrm{R}^T: x = (x_t,\ t \in T),\ x_t = x_{t+}(0 \leqslant t < 1),\ x_{t-}\ \text{存在于}(0,1]\ \text{上}\}.$$

在 D 上引入 Skorohod 尺度 d：

$$\begin{aligned}d(x,y) = \inf\{\varepsilon > 0: \exists\ \lambda \in \Lambda,\ \text{于}\ \sup_t|x_t - y_{\lambda(t)}| \\ + \sup_t|t - \lambda(t)| \leqslant \varepsilon\},\end{aligned}$$

其中

$$\begin{aligned}\Lambda = \{\lambda: \lambda = (\lambda(t),\ t \in T)\ \text{为}\ T\ \text{上连续的严格增函数}, \\ \text{且}\ \lambda(0) = 0,\ \lambda(1) = 1\}.\end{aligned}$$

假如$\mathscr{B}(D)$是由 D 上的柱集类产生的 σ-代数,$\mathscr{B}_0(D)$是由 D 上关于尺度 d 的开集类产生的 σ-代数,则 $\mathscr{B}(D)=\mathscr{B}_0(D)$. 这个结论的证明可以参考文献[4].

(七) 一般乘积空间$\underset{t\in T}{\times}(\Omega_t, \mathscr{F}_t)$

设 T 为任一有序指标集,$(\Omega_t, \mathscr{F}_t)$为可测空间($\forall\ t\in T$). 令

$$\Omega=\underset{t\in T}{\times}\Omega_t=\{\omega:\ \omega=(\omega_t,\ t\in T),\ \omega_t\in\Omega_t(t\in T)\},$$

$$\mathscr{I}^T=\{A:\ A=\mathscr{I}^T_{t_1\cdots t_n}(\underset{i=1}{\overset{n}{\times}}B_i),\ \text{其中}\ t_i\in T,\ t_1<t_2<\cdots<t_n,$$

$$B_i\in\mathscr{F}_{t_i},\ 1\leqslant i<\infty\},$$

$$\mathscr{F}=\underset{t\in T}{\times}\mathscr{F}_t=\sigma(\mathscr{I}^T).$$

则$(\Omega, \mathscr{F})$为可测空间,有时记为$(\Omega, \mathscr{F})=\underset{t\in T}{\times}(\Omega_t, \mathscr{F}_t)$. 前面讨论的所有可测空间均可视作这里的特例. 在某些特别的情形下,这类可测空间可以通过在 Ω 上引入某种拓扑结构产生,如(一)~(六)中的可测空间. 然而,在一般情形下,讨论这两种构作方式之间的关系,则是另一个专门研究的课题.

§1.3 测度的存在与唯一性

定义 1.3.1 设 μ 为集类 $\mathscr{A}$ 上的集函数. 若

(i) $\varnothing\in\mathscr{A}\Rightarrow\mu(\varnothing)=0$;

(ii) $\mu(A)\in[0,\infty]$($\forall\ A\in\mathscr{A}$);

(iii) $\mu(\bigcup\limits_{k=1}^{n}E_k)=\sum\limits_{k=1}^{n}\mu(E_k)$,其中

$$(E_k,\ 1\leqslant k\leqslant n)\subset\mathscr{A}\ \text{不相联,且}\ \bigcup_{k=1}^{n}E_k\in\mathscr{A};$$

则称 μ 为 $\mathscr{A}$ 上的可加测度.

如果条件(iii)改为

(iii)′ $\mu(\bigcup\limits_{k=1}^{\infty}E_k)=\sum\limits_{k=1}^{\infty}\mu(E_k)$,其中

$$(E_k,\ k\geqslant 1)\subset\mathscr{A}\ \text{不相联,且}\ \bigcup_{k=1}^{\infty}E_k\in\mathscr{A};$$

则称 μ 为 $\mathscr{A}$ 上的测度.

若对 $\forall\ A\in\mathscr{A}$,有 $\mu(A)<\infty$,则称 μ 在 $\mathscr{A}$ 上有限.

若对 $\forall\ A\in\mathscr{A}$,存在 $(E_k,\ k\geqslant 1)\subset\mathscr{A}$,使得

$$A\subset\bigcup_{k}E_k,\ \text{且}\ \mu(E_k)<\infty\ (\forall\ k\geqslant 1),$$

则称 μ 在 $\mathscr{A}$ 上 σ-有限.

定理 1.3.1 设 $\mathscr{A}$ 为半环,μ 为 $\mathscr{A}$ 上的可加测度,则 μ 具有下列性质:

(i) 单调性: $A,B\in\mathscr{A}$,且 $A\subset B\Rightarrow\mu(B)\geqslant\mu(A)$.

(ii) 可减性: $A,B\in\mathscr{A}$, $A\subset B$,且 $B\backslash A\in\mathscr{A}$,$\mu(A)<\infty$

$\Rightarrow\mu(B\backslash A)=\mu(B)-\mu(A)$.

(iii) 半可加性: $\mu(A)\leqslant\sum\limits_{k\geqslant 1}^{n}\mu(E_k)$,其中

$$A\in\mathscr{A},\ (E_k,\ 1\leqslant k\leqslant n)\subset\mathscr{A},\ A\subset\bigcup_{k=1}^{n}E_k.$$

(iv) $\sum\limits_{k=1}^{\infty}\mu(E_k)\leqslant\mu(E)$,其中

$$(E,\ E_k,\ k\geqslant 1)\subset\mathscr{A},\ (E_k,\ k\geqslant 1)\ \text{不相联},\ \bigcup_{k=1}^{\infty}E_k\subset E.$$

(其证明留给读者作为练习.)

定理 1.3.2 设 $\mathscr{A}$ 为半环,μ 为 $\mathscr{A}$ 上的可加测度,则 μ 为 $\mathscr{A}$ 上的测度之充要条件是:μ 具有可数半可加性,即

$$\mu(A)\leqslant\sum_{k=1}^{\infty}\mu(E_k),\ \text{其中}\ A,E_k\in\mathscr{A},\ A\subset\bigcup_{k=1}^{\infty}E_k.$$

证 这是定理 1.3.1 的推论.

定理 1.3.3 设 $\mathscr{A}$ 为代数或环,μ 为 $\mathscr{A}$ 上的可加测度,且有限,则 μ 为 $\mathscr{A}$ 上测度的充要条件是:μ 具有上半连续性,即

$\mu(A_n)\downarrow\mu(A)$(或 0),其中

$(A,\ A_n,\ n\geqslant)\subset\mathscr{A}$,且 $A_n\downarrow A$(或$\varnothing$).

证 令 $A_n'=A_n\backslash A$,则 $A_n'\in\mathscr{A}$,且 $A_n'\downarrow\varnothing$. 因此,不妨假

定：$A_n\downarrow\varnothing$. 现在，定义 $B_n=A_1\backslash A_n$，则 $B_n\uparrow A_1$. 令 $B_1'=B_1$，$B_n'=B_n\backslash\bigcup_{k=1}^{n-1}B_k(n\geqslant 2)$，则$(B'_n,\ n\geqslant 1)\subset\mathscr{A}$，且不相联，$B_n=\bigcup_{k=1}^{n}B'_k$. 按 μ 的可加性，$\mu(\bigcup_{k=1}^{n}B_k')=\sum_{k=1}^{n}\mu(B_k')$，按 μ 有限，$\mu(B_n)\leqslant\mu(A_1)<\infty$. 于是，$\sum_k\mu(B'_k)$收敛.

$$\Rightarrow\lim_{n\to\infty}\mu(B_n)=\sum_k\mu(B_k').$$

若 μ 为测度，则 $\sum_k\mu(B_k')=\mu(\bigcup_k B_k')=\mu(A_1)$.

$$\Rightarrow\lim_{n\to\infty}\mu(A_n)=0.$$

若$\lim_{n\to\infty}\mu(A_n)=0$，则 $\mu(A_1)=\lim_{n\to\infty}\mu(B_n)=\sum_k\mu(B_k')$

$$\Rightarrow\mu(\bigcup_k B_k')=\sum_k\mu(B_k').$$

$\Rightarrow\mu$ 具有 σ-可加性，即 μ 为测度.

注 1 在上述定理中，"μ 有限"之条件不可少. 例如，设 $\Omega=(0,1]$，$\mathscr{A}$ 为 Ω 上的子集之全体. 令

$$\mu(A)=\begin{cases}+\infty, & A\neq\varnothing,\\ 0, & A=\varnothing.\end{cases}$$

则 μ 为 $\mathscr{A}$ 上的测度. 但不具有上半连续性.

定理 1.3.4 设 μ 为代数 $\mathscr{A}$ 上的 σ-有限测度，则 $\mathscr{A}$ 不能包含一个不可数、不相联，且具有正 μ-测度的集族.

证 设 Θ 为具有如下性质的参数集：

$$B_\theta\in\mathscr{A}\ (\forall\ \theta\in\Theta),\quad \mu(B_\theta)>0(\forall\ \theta\in\Theta),$$

$$B_{\theta_1}\cap B_{\theta_2}=\varnothing(\theta_1、\theta_2\in\Theta,\ \theta_1\neq\theta_2).$$

则此定理等价于证明：Θ 至多为一个可数集.

设 $A\in\mathscr{A}$，且 $\mu(A)<\infty$. 若$(\theta_k,\ 1\leqslant k\leqslant n)\subset\Theta$，且不相同，并满足 $\mu(A\cap B_{\theta_k})\geqslant\varepsilon>0$，则 $n\varepsilon\leqslant\sum_{k=1}^{n}\mu(A\cap B_{\theta_k})\leqslant\mu(A)$. 因此，对$\forall\ \varepsilon>0$，集$\{\theta\in\Theta:\ \mu(A\cap B_\theta)\geqslant\varepsilon\}$只能含有限个元素. 按 μ 的 σ-有限

性，存在$A_i\in\mathscr{A}$，使得$\Omega=\bigcup_{i=1}^{\infty}A_i$，且$\mu(A_i)<\infty\ (\forall\ i\geqslant 1)$. 令$\Theta_k=\{\theta\in\Theta:\ \mu(A_k\cap B_\theta)>0\}$，则$\Theta_k$至多可数. 若$\theta\notin\Theta_k$，则$\mu(A_k\cap B_\theta)=0$. 于是，当$\theta\notin\bigcup_{i=1}^{\infty}\Theta_k$时，

$$\mu(B_\theta)\leqslant\sum_k\mu(A_k\cap B_\theta)=0.\ \Rightarrow\theta\notin\Theta\Rightarrow\Theta=\bigcup_k\Theta_k.$$

即Θ至多可数.

定理 1.3.5 设$\mathscr{P}$为Ω上的π-类. μ_1,μ_2为$\sigma(\mathscr{P})$上的测度，且在$\mathscr{P}$上σ-有限. Ω具有$\mathscr{P}$-集覆盖. 若

$$\mu_1(A)=\mu_2(A)\ (\forall\ A\in\mathscr{P}),$$

则

$$\mu_1(A)=\mu_2(A)\ (\forall\ A\in\sigma(\mathscr{P})).$$

证 令

$$\mathscr{L}_B=\{A:\ A\in\sigma(\mathscr{P}),\text{且}\ \mu_1(B\cap A)=\mu_2(B\cap A)\}.$$

则当$B\in\mathscr{P}$，$\mu_1(B)<\infty$时，$\mathscr{L}_B$为λ-类.

事实上，按假设，$\Omega\in\mathscr{L}_B$，假定$B\in\mathscr{L}_B$，则

$$\begin{aligned}\mu_1(B\cap A^c)&=\mu_1(B\backslash A\cap B)=\mu_1(B)-\mu_1(A\cap B)\\&=\mu_2(B)-\mu_2(A\cap B)=\mu_2(B\cap A^c).\end{aligned}$$

$\Rightarrow A^c\in\mathscr{L}_B$. 按测度的连续性，$\mathscr{L}_B$为单调类. 综上分析，便可得结论：$\mathscr{L}_B$为$\lambda$-类.

显然，对$\forall\ B\in\mathscr{P}$，$\mu_1(B)<\infty$，有$\mathscr{P}\subset\mathscr{L}_B$. 按定理 1.1.1，$\sigma(\mathscr{P})\subset\mathscr{L}_B$. 现在，利用这一事实来证明，对$\forall\ A\in\sigma(\mathscr{P})$，有$\mu_1(A)=\mu_2(A)$. 事实上，按$\mu_1,\mu_2$在$\mathscr{P}$上$\sigma$-有限及$\Omega$具有$\mathscr{P}$-集覆盖，$\exists\ \mathscr{P}$-集列$(B_k,\ k\geqslant 1)$，$\ni\ \Omega=\bigcup_{k=1}^{\infty}B_k$，且$\mu_2(B_k)=\mu_1(B_k)<\infty\ (\forall\ k\geqslant 1)$. 令$B_1'=B_1$，$B_n'=B_n\backslash\bigcup_{k=1}^{n-1}B_k\ (n\geqslant 2)$. 则$\Omega=\bigcup_{k=1}^{\infty}B_k'$，且$(B_k',\ k\geqslant 1)$不相联，$\mu_i(B_k')<\infty\ (i=1,2.\ k\geqslant 1)$. 显然，$\mu_1(B_1')=\mu_2(B_1')$. 假定$n\geqslant 2$，按$\sigma(\mathscr{P})\subset\mathscr{L}_{B_n}$，有

$$\mu_1(B_n')=\mu_1(B_n\backslash\bigcup_{k=1}^{n-1}B_k)=\mu_1(B_n\backslash\bigcup_{k=1}^{n-1}B_k')$$

$$=\mu_1(B_n)-\mu_1(\bigcup_{k=1}^{n-1}B'_k\cap B_n)$$

$$=\mu_1(B_n)-\sum_{k=1}^{n-1}\mu_1(B'_k\cap B_n)$$

$$=\mu_2(B_n)-\sum_{k=1}^{n-1}\mu_2(B'_k\cap B_n)=\mu_2(B'_n).$$

同理可证：$\mu_1(B'_n\cap A)=\mu_2(B'_n\cap A)\ (\forall\ A\in\sigma(\mathscr{P}))$.

$$\Rightarrow\ \mu_1(A)=\mu_1(\bigcup_{k=1}^{\infty}B'_k\cap A)=\sum_k\mu_1(B'_k\cap A)$$

$$=\sum_k\mu_2(B'_k\cap A)=\mu_2(A)\ (\forall\ A\in\sigma(\mathscr{P})).$$

注 2 上述定理中"Ω 具有 $\mathscr{P}$-集覆盖"之条件不能少. 例如，设 $\mathscr{P}=\{\varnothing\}$，$\sigma(\mathscr{P})=\{\varnothing,\Omega\}$，$\mu_1(\Omega)\neq\mu_2(\Omega)$，$\mu_1(\varnothing)=0=\mu_2(\varnothing)$. 则上述定理对这种情形就不成立. 原因在于 Ω 不具有 $\mathscr{P}$-集覆盖.

定义 1.3.2 设 μ^* 为 Ω 的子集之全体上的集函数，且具有下列性质：

(i) $\mu^*(A)\in[0,\infty]\ (\forall\ A\subset\Omega)$；

(ii) $\mu^*(\varnothing)=0$；

(iii) μ^* 具有单调性，即 $\mu^*(A)\leqslant\mu^*(B)\ (A\subset B\subset\Omega)$；

(iv) μ^* 具有可数半可加性，即 $\mu^*(\bigcup_n A_n)\leqslant\sum_n\mu^*(A_n)$.

则称 μ^* 为 Ω 上的一个外测度.

定理 1.3.6 设 $\mathscr{A}$ 为 Ω 上的一个子集类. $\varnothing\in\mathscr{A}$，$\rho$ 为 $\mathscr{A}$ 上的集函数，$\rho(\varnothing)=0$，$\rho(A)\in[0,\infty]\ (\forall\ A\in\mathscr{A})$. 令

$$\mu^*(E)=\begin{cases}\inf\{\sum_n\rho(A_n):(A_n,\ n\geqslant 1)\text{ 为 }E\text{ 的 }\mathscr{A}\text{- 集覆盖}\},\\ \infty,\quad\text{若不存在 }E\text{ 的 }\mathscr{A}\text{- 集覆盖}.\end{cases}$$

则 μ 为 Ω 上的一个外测度(常称此测度由 ρ 产生).

证 易证定义 1.3.2 中的前三条. 下证 μ^* 满足条件(iv)

假如 $E=\bigcup_n A_n$ 不存在 $\mathscr{A}$-集覆盖，则至少有一个，比如 A_{n_0}，

不存在 $\mathscr{A}$-集覆盖. 这时 $\mu^*(A_{n_0})=\infty$. 从而,

$$\mu^*(E)\leqslant\infty=\mu^*(A_{n_0})\leqslant\sum_n\mu^*(A_n)\Rightarrow\text{(iv) 成立}.$$

现假定 E 具有 $\mathscr{A}$-集覆盖,则每个 A_n 都具有 $\mathscr{A}$-集覆盖. 于是,对$\forall\ \varepsilon>0$ 及 $n\geqslant1$,$\exists\ (A_{nk},\ k\geqslant1)\subset\mathscr{A}$,使得

$$A_n\subset\bigcup_{k=1}^{\infty}A_{nk},\quad\sum_k\rho(A_{nk})\leqslant\mu^*(A_n)+\frac{1}{2^n}\varepsilon.$$

$$\Rightarrow\mu^*(\bigcup_n A_n)\leqslant\sum_n\sum_k\rho(A_{nk})\leqslant\sum_n\mu^*(A_n)+\varepsilon.$$

$\Rightarrow$(iv)成立.

定义 1.3.3 设 μ^* 为 Ω 上的非负集函数. 若

$$\mu^*(A\cap E)+\mu^*(A^c\cap E)=\mu^*(E)\ (\forall\ E\subset\Omega).\tag{1.3-1}$$

则称 A 为 μ^*-可测集. 符号 $\mathscr{M}(\mu^*)$表示 μ^*-可测集全体.

注 3 若 μ^* 为 Ω 上的一个外测度,则条件(1.3-1)等价于

$$\mu^*(A\cap E)+\mu^*(A^c\cap E)\leqslant\mu^*(E)\ (\forall\ E\subset\Omega).\tag{1.3-2}$$

定理 1.3.7 若 μ^* 为 Ω 上的一个外测度,则 $\mathscr{M}(\mu^*)$为 σ-代数,而 μ^* 到 $\mathscr{M}(\mu^*)$上的限制为 $\mathscr{M}(\mu^*)$上的测度.

证 首先证明:$\mathscr{M}(\mu^*)$为代数.

显然,$\varnothing,\Omega\in\mathscr{M}(\mu^*)$. 按对称性,$A^c\in\mathscr{M}(\mu^*)$当 $A\in\mathscr{M}(\mu^*)$时. 设 $A,B\in\mathscr{M}(\mu^*)$,$E\subset\Omega$,按(1.3-1)式,有

$$\mu^*(E)=\mu^*(B\cap E)+\mu^*(B^c\cap E).$$
$$\mu^*(BE)=\mu^*(ABE)+\mu^*(A^cBE).$$
$$\mu^*(B^cE)=\mu^*(AB^cE)+\mu^*(A^cB^cE).$$
$$A^cBE\cup AB^cE\cup A^cB^cE=(AB)^cE.$$
$$\Rightarrow\mu^*(E)\geqslant\mu^*(ABE)+\mu^*((AB)^cE)$$
$$\Rightarrow AB\in\mathscr{M}(\mu^*)\ \text{(按(1.3-2) 式)}.$$

其次证明:μ^* 在 $\mathscr{M}(\mu^*)$上具有可加性. 事实上,对 $A,B\in\mathscr{M}(\mu^*)$,$AB=\varnothing$,有

$$\mu^*(A\cup B)=\mu^*(A(A\cup B))+\mu^*(A^c(A\cup B))$$
$$=\mu^*(A)+\mu^*(B).$$

最后证明：$\mathscr{M}(\mu^*)$为σ-代数，且μ^*为$\mathscr{M}(\mu^*)$上的测度. 事实上，设$(A_n, n\geqslant 1)\subset\mathscr{M}(\mu^*)$不相联. $A=\bigcup_n A_n$. 按μ^*的可加性，对$\forall\ m\geqslant 1$，有

$$\sum_{n\leqslant m}\mu^*(A_n)=\mu^*(\bigcup_{n\leqslant m}A_n)\leqslant\mu^*(A)$$
$$\Rightarrow\sum_n\mu^*(A_n)\leqslant\mu^*(A).$$

由μ^*具有可数半可加性知：

$$\mu^*(A)=\sum_n\mu^*(A_n).$$

因此，“μ^*为$\mathscr{M}(\mu^*)$上的测度”等价于“$A\in\mathscr{M}(\mu^*)$”(即$\mathscr{M}(\mu^*)$为σ-代数). 令$B_m=\bigcup_{n\leqslant m}A_n$，则$B_m\in\mathscr{M}(\mu^*)$.

$$\Rightarrow\mu^*(EB_m)=\mu^*(B_{m-1}EB_m)+\mu^*(B_{m-1}^cEB_m)$$
$$=\mu^*(B_{m-1}E)+\mu^*(EA_m).\ (\forall\ E\subset\Omega).$$

由此递推关系得

$$\mu^*(EB_m)=\sum_{n=1}^m\mu^*(EA_n)\ (m\geqslant 1,\ E\subset\Omega).$$

$$\Rightarrow\mu^*(E)=\mu^*(B_mE)+\mu^*(B_m^cE)$$
$$\geqslant\sum_{k=1}^m\mu^*(EA_k)+\mu^*(A^cE)\ (\forall\ E\subset\Omega,\ m\geqslant 1)$$
$$\Rightarrow\mu^*(E)\geqslant\sum_k\mu^*(EA_k)+\mu^*(A^cE)$$
$$\geqslant\mu^*(EA)+\mu^*(A^cE).$$

故按(1.3-2)式，$A\in\mathscr{M}(\mu^*)$.

推论 1.3.1 设$\mathscr{A}$为Ω上的半环. ρ为$\mathscr{A}$上的测度. 则由ρ产生的外测度μ^*为$\sigma(\mathscr{A})$上的测度，且

$$\mu^*(A)=\rho(A)\ (\forall\ A\in\mathscr{A}).$$

证 假如$\mathscr{A}\subset\mathscr{M}(\mu^*)$，则按定理 1.3.7，$\sigma(\mathscr{A})\subset\mathscr{M}(\mu^*)$.

从而，μ^* 为 $\sigma(\mathscr{A})$上的测度. 按定理 1.3.6 及 ρ 为 $\mathscr{A}$ 上的测度，立即可得 $\mu^*(A)=\rho(A)$ $(\forall\ A\in\mathscr{A})$. 剩下的问题就是证明：$\mathscr{A}\subset\mathscr{M}(\mu^*)$.

设 $A\in\mathscr{A}$，$E\subset\Omega$，若 E 不存在 $\mathscr{A}$-集覆盖，则

$$\mu^*(E)=\infty\Rightarrow\mu^*(AE)+\mu^*(A^cE)\leqslant\infty=\mu^*(E).$$

否则，对$\forall\ \varepsilon>0$，$\exists\ E$ 的 $\mathscr{A}$-集覆盖$(A_n,\ n\geqslant1)$，$\ni E\subset\bigcup\limits_n A_n$，且

$$\sum_n\rho(A_n)\leqslant\mu^*(E)+\varepsilon.$$

令 $B_n=AA_n$，则 $B_n\in\mathscr{A}$，且 $A^cA_n=A_n\backslash B_n=\bigcup\limits_{k=1}^{m_n}C_{nk}(C_{nk}\in\mathscr{A})$.

$$\begin{aligned}\Rightarrow\mu^*(AE)+\mu^*(A^cE)&\leqslant\sum_n\rho(AA_n)+\sum_n\sum_{k=1}^{m_n}\rho(C_{nk})\\&=\sum_n\Big(\rho(B_n)+\sum_{k=1}^{m_n}\rho(C_{nk})\Big)\\&=\sum_n\rho(A_n)\leqslant\mu^*(E)+\varepsilon.\end{aligned}$$

$\Rightarrow A\in\mathscr{M}(\mu^*)$ （按(1-3-2)式），即 $\mathscr{A}\subset\mathscr{M}(\mu^*)$.

推论 1.3.2 设 $\mathscr{A}$ 为 Ω 上的半环. Ω 具有 $\mathscr{A}$-集覆盖. ρ 为 $\mathscr{A}$ 上的 σ-有限测度. 则 ρ 具有到 $\sigma(\mathscr{A})$上的唯一 σ-有限测度延拓，即存在 $\sigma(\mathscr{A})$上的唯一 σ-有限测度 ρ^*，使得

$$\rho^*(A)=\rho(A)\ (\forall\ A\in\mathscr{A}).$$

证 按推论 1.3.1，ρ^* 存在. 按定理 1.3.5，ρ^* 唯一.

定理 1.3.8 设 $\mathscr{A}$ 为 Ω 上的半环. Ω 具有 $\mathscr{A}$-集覆盖. μ 为 $\mathscr{F}=\sigma(\mathscr{A})$上的测度，且在 $\mathscr{A}$ 上 σ-有限，则

(1) 对$\forall\ B\in\mathscr{F}$，$\varepsilon>0$，存在不相联的 $\mathscr{A}$-集列$(A_k,\ k\geqslant1)$，使得

$$B\subset\bigcup_k A_k,\quad \mu\Big(\bigcup_k A_k\backslash B\Big)<\varepsilon.$$

(2) 对$\forall\ B\in\mathscr{F}$，$\varepsilon>0$，$\mu(B)<\infty$，存在有限个不相联的 $\mathscr{A}$-集列$(A_k,\ 1\leqslant k\leqslant n)$，$\ni\mu\big(B\triangle(\bigcup\limits_{k=1}^{n}A_k)\big)<\varepsilon$ （"$\triangle$"表示对称差）.

证 假如 μ^* 是由 μ 在 $\mathscr{A}$ 上的值定义的外测度，则按推论 1.3.2，有 $\mu^*=\mu$ 于 $\sigma(\mathscr{A})$ 上，假定 $B\in\mathscr{F}$，且 $\mu(B)<\infty$，则 $\mu^*(B)=\mu(B)<\infty$. 从而，$\exists$ $\mathscr{A}$-集列 $(A_k,\ k\geqslant 1)$，使得

$$B\subset\bigcup_k A_k,\ \mu(\bigcup_k A_k)\leqslant\sum_n\mu(A_n)\leqslant\mu(B)+\varepsilon.$$

$$\Rightarrow\mu(\bigcup_k A_k\setminus B)<\varepsilon.$$

令 $A'_1=A_1$，

$$A'_n=A_n\setminus\bigcup_{k=1}^{n-1}A_k=A_nA_1^c\cdots A_{n-1}^c\ (n\geqslant 2).$$

则 $(A'_n,\ n\geqslant 1)$ 不相联，且由 $\mathscr{A}$ 为半环可知：$A'_n=\bigcup_{k=1}^{m_n}C_{nk}$（利用归纳法即可得这一表示）. 重排之后，得

$$\bigcup_k A_k=\bigcup_k A'_k=\bigcup_k C_k,$$

其中 $C_k\in\mathscr{A}\ (k\geqslant 1)$，且 $(C_k,\ k\geqslant 1)$ 不相联.

$$\Rightarrow\mu(\bigcup_k C_k\setminus B)=\mu(\bigcup_k A_k\setminus B)\leqslant\varepsilon.$$

令 $A=\bigcup_k C_k$，则由测度的连续性，$\exists$ $n\geqslant 1$，$\ni$ $\mu(A\setminus\bigcup_{k=1}^n C_k)\leqslant\varepsilon$. 故

$$\mu(B\triangle(\bigcup_{k=1}^n C_k))=\mu(B\setminus\bigcup_{k=1}^n C_k)+\mu(\bigcup_{k=1}^n C_k\setminus B)$$

$$\leqslant\mu(A\setminus\bigcup_{k=1}^n C_k)+\mu(A\setminus B)\leqslant 2\varepsilon.$$

上述分析表明：当 $B\in\mathscr{F}$，$\mu(B)<\infty$ 时，定理真.

现设 $B\in\mathscr{F}$，且 $\mu(B)=\infty$. 按定理的条件，$\exists$ $\mathscr{A}$-集列 $(C_m,\ m\geqslant 1)$，$\ni$ $\Omega=\bigcup_m C_m$，且 $\mu(C_m)<\infty\ (\forall\ m\geqslant 1)$. 因此，对每个 $B\cap C_m$，结论(1)成立，即$\exists$ $\mathscr{A}$-集列 $(A_{mk},\ k\geqslant 1)$，使得

$$B\cap C_m\subset\bigcup_k A_{mk},\quad \mu(\bigcup_k A_{mk}\setminus B\cap C_m)<\frac{\varepsilon}{2^m}.$$

$$\Rightarrow\mu(\bigcup_{m,k}A_{mk}\setminus B)\leqslant\mu(\bigcup_m(\bigcup_k A_{mk}\setminus B\cap C_m))\leqslant\sum_{m=1}^{\infty}\frac{\varepsilon}{2^m}=\varepsilon.$$

$\Rightarrow$结论(1) 成立

推论 1.3.3 设 $\mathscr{A}$ 为 Ω 上的代数. $\mathscr{F}=\sigma(\mathscr{A})$. μ 为 $\mathscr{F}$ 上的

有限测度. 则对$\forall\ B\in\mathscr{F}$, $\varepsilon>0$, $\exists\ A\in\mathscr{A}$,

$$\ni\ \mu(A\triangle B)<\varepsilon.$$

证 事实上, 代数必为半环, 且 $\Omega\in\mathscr{A}$, 又 μ 为有限测度必 σ-有限. 故按定理 1.3.7(2)即得所要的结论.

定义 1.3.4(Caratheodory 条件). 设 μ^* 为 R^k 上的一个外测度. 记

$$\mathrm{dist}(A,B)=\inf\{|x-y|:x\in A,\ y\in B\}.$$

如果 $\mathrm{dist}(A,B)>0$, 必有

$$\mu^*(A\cup B)=\mu^*(A)+\mu^*(B).$$

则说 μ^* 满足 Caratheodory 条件, 简称 μ^* 满足条件(C).

定理 1.3.9 若 μ^* 为 R^k 上的一个外测度, 且满足条件(C), 则

$$\sigma(\{A:A=(a,b],\ a,b\in\mathrm{R}^k\})\subset\mathscr{M}(\mu^*).$$

证 令 $\mathscr{I}_k=\{(a,b]:a,b\in\mathrm{R}^k\}$, $\mathscr{L}_k$ 为 R^k 中闭子集全体. 则 $\sigma(\mathscr{I}_k)=\sigma(\mathscr{L}_k)=\mathscr{B}^k$. 按定理 1.3.7, 仅需证明: $\mathscr{L}_k\subset\mathscr{M}(\mu^*)$. 设

$$A\in\mathscr{L}_k,\quad E\subset\mathrm{R}^k.\ \text{记}\ B'=AE,\ B''=A^cE,$$

$$C_n=\left\{x\in B'':\ \mathrm{dist}(x,A)\geqslant\frac{1}{n}\right\}.$$

按条件(C), 有

$$\mu^*(E)=\mu^*(B'\cup B'')\geqslant\mu^*(B'\cup C_n)=\mu^*(B')+\mu^*(C_n).$$

若 $\mu^*(B'')\leqslant\lim\limits_{n\to\infty}\mu^*(C_n)$, 则 $A\in\mathscr{M}(\mu^*)$. 从而, $\mathscr{L}_k\subset\mathscr{M}(\mu^*)$.

令 $D_n=C_{n+1}\backslash C_n$, 则当 $x\in D_{n+1}$, $|x-y|<\dfrac{1}{n(n+1)}$ 时,

$$\mathrm{dist}(y,A)\leqslant|y-x|+\mathrm{dist}(x,A)$$

$$<\frac{1}{n(n+1)}+\frac{1}{n+1}=\frac{1}{n}.$$

$\Rightarrow y\notin C_n$.

$\Rightarrow\mathrm{dist}(D_{n+1},C_n)=\inf\{|x-y|:x\in D_{n+1},\ y\in C_n\}$

$$\geqslant\frac{1}{n(n+1)}.$$

这里假定 D_{n+1}，C_n 非空. 由此即知：当 D_{2n}非空时，$(D_{2k}, 1\leqslant k\leqslant n)$ 相互之间有正距离. 按条件(c)，有

$$\mu^*(C_{2n}) = \mu^*(\bigcup_{k=1}^{n} D_{2k}) = \sum_{k=1}^{n}\mu^*(D_{2k}).$$

$$\mu^*(C_{2n+1}) \geqslant \mu^*(C_{2n}) = \sum_{k=1}^{n}\mu^*(D_{2k}). \tag{1.3-3}$$

同理可知：当 D_{2n-1}非空时，$\{D_1, D_3, \cdots, D_{2n-1}\}$相互之间有正距离($D_1=C_1$)，且

$$\mu^*(C_{2n}) \geqslant \mu^*(C_{2n-1}) = \sum_{k=1}^{n}\mu^*(D_{2k-1}). \tag{1.3-4}$$

按 μ^* 的半可加性，有

$$\mu^*(B'') = \mu^*(\bigcup_k C_k) \leqslant \mu^*(C_{2n}) + \sum_{k=n}^{\infty}\mu^*(D_{2k}) + \sum_{k=n+1}^{\infty}\mu^*(D_{2k-1}). \tag{1.3-5}$$

如果级数 $\sum_k \mu^*(D_{2k})$ 或 $\sum_k \mu^*(D_{2k-1})$ 发散，则按(1.3-3)式或(1.3-4)式，立即可得

$$\mu^*(B'') \leqslant \infty = \lim_{n\to\infty}\mu^*(C_n).$$

如果两个级数都收敛，则按(1.3-5)式，有

$$\mu^*(B'') \leqslant \lim_{n\to\infty}\mu^*(C_n).$$

在测度论中，Euclide 空间上的 Lebesque 测度是最典型的测度例. 考虑 k-维 Borel 可测空间$(\mathrm{R}^k, \mathscr{B}^k)$. $\mathscr{A}_0^k=\{(a,b]: a,b\in \mathrm{R}^k\}$. 这是 R^k 上的半环，且 $\Omega=\mathrm{R}^k$ 具有 $\mathscr{A}_0^k$-集覆盖，定义 $\mathscr{A}_0^k$ 上的集函数 λ_k：

$$\lambda_k((a,b]) = \prod_{i=1}^{k}(b_i - a_i) \quad (\forall\ (a,b]\in\mathscr{A}_0^k, \text{且}\ a\leqslant b).$$

显然，λ_k 为 $\mathscr{A}_0^k$ 上的 σ-有限测度. 按唯一性定理 1.3.5 及推论 1.3.2，λ_k 有到 $\mathscr{B}^k$ 上的唯一延拓，仍记为 λ_k，称 λ_k 为$(\mathrm{R}^k, \mathscr{B}^k)$上的 Lebesque 测度. 这种测度的一个最基本的性质是平移不变性，即

$\lambda_k(A)=\lambda_k(A+x)(x\in\mathrm{R}^k,\ A\in\mathscr{B}^k)$. 另一个重要特性就是正则性. 例如 μ 为 $\mathscr{B}^k$ 上的测度, 且对 $\mathscr{B}^k$ 中的任意有界元 A, 有 $\mu(A)<\infty$, 则该 μ 为 $\mathscr{B}^k$ 上的正则测度. 显然, λ_k 是正则的. 正则测度 μ 具有下列重要性质: 对 $\forall\ A\in\mathscr{B}^k$ 及 $\varepsilon>0$, 存在闭集 C 和开集 G, 使 $C\subset A\subset G$, $\mu(G\backslash C)<\varepsilon$; 若 $\mu(A)<\infty$, 则

$$\mu(A) = \sup\{\mu(K):\ K \text{ 为 } A \text{ 中的紧子集}\}.$$

注 3 一个外测度 μ^* 要能成为 $\sigma(\mathscr{A})$ 上的测度除去 $\mu^*|_{\mathscr{A}}$ 为 $\mathscr{A}$ 上的测度外, 必须有 $\mathscr{A}\subset\mathscr{M}(\mu^*)$. 条件(C)正是保证了这一点. 然而, 若 $\mathscr{A}$ 只是一个 π-类, 则一般不能保证 $\mathscr{A}\subset\mathscr{M}(\mu^*)$. 例如, 设 $\Omega=(0,1]$, $\lambda(0,b]=b$, $\mathscr{A}=\{(0,b]:\ b\in[0,1]\}$, μ^* 是由 λ 产生的外测度, 则易证:

$$\mu^*(A) = \max\{x:\ x\in A\}\ (\forall\ A\in\Omega),$$

$$\mathscr{M}(\mu^*) = \{\varnothing,\Omega\}.$$

可见, $\mathscr{A}\not\subset\mathscr{M}(\mu^*)$. 在这种情形下, 显然, λ 为 $\mathscr{A}$ 上的测度, 且 $\lambda|_{\mathscr{A}}=\mu^*|_{\mathscr{A}}$. 但 μ^* 不为 $\sigma(\mathscr{A})$ 上的测度. 原因在于 $\mathscr{A}$ 仅为 π-类, 而不是半环.

§ 1.4 可测空间上的概率测度

上节讨论可测空间上的一般测度. 本节讨论的问题是在某些常用可测空间上如何引入概率测度. 如果 P 为可测空间 $(\Omega,\mathscr{F})$ 上的测度, 且 $P(\Omega)=1$, 则称 P 为概率测度.

(一) 可测空间 $(\mathrm{R},\mathscr{B})$ 上的概率测度

设 P 为 $(\mathrm{R},\mathscr{B})$ 上的概率测度, 则可定义 R 上的一实值函数:

$$F(x) = P((-\infty,x])\ (\forall\ x\in\mathrm{R}). \tag{1.4-1}$$

此函数具有下列性质:

(i) $F(-\infty)=\lim\limits_{x\to-\infty}F(x)$, $F(+\infty)=\lim\limits_{x\to\infty}F(x)=1$.

(ii) $F(x)$, $x\in\mathrm{R}$, 是单调不降的.

(iii) F 右连左极，即 $F(x)=F(x+)$，$\lim\limits_{u\uparrow x}F(u)$存在($x\in \mathrm{R}$).

定义 1.4.1 若映射 F: R→R 具有上述三条性质，则称 F 为 R 上的分布函数.

现在考虑一个反问题，假如 F 为 R 上的一个分布函数. 试问能否由此分布函数唯一地产生$(\mathrm{R},\mathscr{B})$上的概率测度 P，使得(1.4-1)式成立.

定理1.4.1 设 F 为R 上的分布函数，则存在$(\mathrm{R},\mathscr{B})$上的唯一概率测度 P，使得

$$P((a,b]) = F(b) - F(a).\ (\forall\ a,b \in \mathrm{R},\ a \leqslant b).$$

证 令 $\mathscr{A}_0=\{(a,b]:\ a,b\in \mathrm{R},\ a\leqslant b\}$，则 $\mathscr{A}_0$ 为 R 上的半环，且 R 具有 $\mathscr{A}_0$-集覆盖，$\mathscr{B}=\sigma(\mathscr{A}_0)$. 定义 $\mathscr{A}_0$ 上的集函数 P_0 如下：

$$P_0((a,b]) = F(b) - F(a)\ (\forall\ a,b \in \mathrm{R},\ a \leqslant b).$$

按推论 1.3-2，若 P_0 为 $\mathscr{A}_0$ 上的测度，则本定理成立. 按定理 1.3.3，P_0 为 $\mathscr{A}_0$ 上的测度等价于：P_0 在 $\mathscr{A}_0$ 上可加，且具有上半连续性.

设 $A=(a_1,\ b_1]$，$B=(a_2,\ b_2]$，$A\cap B=\varnothing$，$A\cup B\in\mathscr{A}_0$，则 $A\cup B=(c_1,\ c_2]$，其中 $c_1=a_1$，$c_2=b_2$ 或 $c_1=a_2$，$c_2=b_1$；$b_1=a_2$ 或 $b_2=a_1$. 这两种情形均有

$$P_0(A\cup B) = P_0(A) + P_0(B).$$

即 P_0 为 $\mathscr{A}_0$ 上的可加集函数.

剩下的问题就是证明：P_0 上半连续. 注意，$\mathscr{A}_0$ 中的任一元素都是有界的，证明的关键在于如下事实：

$$P_0((a',b]) \xrightarrow[a'\downarrow a]{} P_0((a,b])\quad (\text{按 } F \text{ 的右连续性}).$$

设$(A_n,\ n\geqslant 1)\subset\mathscr{A}_0$，且 $A_n\downarrow\varnothing$，则对$\forall\ \varepsilon>0$，$\exists\ B_n=(c_n,b_n]\in\mathscr{A}_0$，∋ $[B_n]\subset A_n$，$A_n\backslash B_n\in\mathscr{A}_0$，且 $P_0(A_n\backslash B_n)<\frac{1}{2^n}\varepsilon$. 令

$$A_1(\delta) = (a_1-\delta,\ b_1+\delta)\quad (\delta>0),$$

则$\{A_1(\delta)\backslash[B_n],\ n\geqslant 1\}$构成紧集$[A_1]$的一个开覆盖，([·]表示闭

包）. 根据有限覆盖定理，

$$\exists\, n_0 \geqslant 1,\ \ni [A_1] \subset \bigcup_{k=1}^{n_0} A_1(\delta)\backslash[B_k].$$

$$\Rightarrow [A_1] = \bigcup_{k=1}^{n_0} [A_1]\backslash[B_k] \Rightarrow \bigcap_{k=1}^{n_0} [B_k] = \varnothing \Rightarrow \bigcap_{k=1}^{n_0} B_k = \varnothing.$$

注意，当 $k \leqslant n_0$ 时，$A_{n_0} \subseteq A_k$，$A_{n_0}\backslash B_k \in \mathscr{A}_0$. 于是，

$$P_0(A_{n_0}) = P_0\Big(A_{n_0}\backslash \bigcap_{k=1}^{n_0} B_k\Big) = P_0\Big(\bigcup_{k=1}^{n_0} A_{n0}\backslash B_k\Big)$$

$$\leqslant \sum_{k=1}^{n_0} P_0(A_{n_0}\backslash B_k) \leqslant \sum_{k=1}^{n_0} P_0(A_k\backslash B_k)$$

$$\leqslant \sum_{k=1}^{n_0} \frac{1}{2^k}\varepsilon < \varepsilon.$$

$$\Rightarrow P_0(A_n) \downarrow 0.$$

推论 1.4.1 设 F 为 R 上的分布函数，则存在 $(\mathrm{R}, \mathscr{B})$ 上的唯一概率测度 P，使得

$$P((-\infty, x]) = F(x) \quad (\forall\, x \in \mathrm{R})$$

证 令 $\mathscr{P} = \{(-\infty, x]: x \in \mathrm{R}\}$，则 $\mathscr{P}$ 为 R 上的一个 π-类. 且 R 具有 $\mathscr{P}$-集覆盖，$\mathscr{B} = \sigma(\mathscr{P})$. 定义

$$P_0((-\infty, x]) = F(x) \quad (\forall\, x \in \mathrm{R}).$$

则易证：P_0 为 $\mathscr{P}$ 上的有限测度. 按定理 1.4.1，$\exists$ $\mathscr{B}$ 上的概率测度 P，$\ni$ $P|_{\mathscr{P}} = P_0$，按定理 1.3.5，此测度 P 是唯一的.

（二）可测空间 $(\mathrm{R}^n, \mathscr{B}^n)$ 上的概率测度

假如 P 为 $(\mathrm{R}^n, \mathscr{B}^n)$ 上的概率测度. 定义函数 F：

$$\mathrm{R}^n \to R:\ F = F(x) = P((-\infty, x]) \ (\forall\, x \in \mathrm{R}^n). \tag{1.4-2}$$

这里 $(-\infty, x] = \mathop{\times}\limits_{k=1}^{n} (-\infty, x_k]$ $(x_k \in \mathrm{R})$，则 F 具有下列性质：

(i) $F(\infty, \cdots, \infty) = \lim\limits_{\substack{x_k \to \infty \\ 1 \leqslant k \leqslant n}} F(x_1, x_2, \cdots, x_n) = 1$，

$$F(y)=0 \text{ 当 } y \text{ 中有一个分量为 } -\infty \text{ 时}.$$

(ii) $F(x)=\lim\limits_{y\downarrow x}F(y)$，$\lim\limits_{y\uparrow x}F(y)$存在，

其中“$y\downarrow x$”$\Longleftrightarrow$“$y_k\downarrow x_k(1\leqslant k\leqslant n)$”.

(iii) $\prod\limits_{k=1}^{n}\Delta_{a_kb_k}F(x)\geqslant 0\ (a,b\in\mathrm{R}^n,\ a\leqslant b)$，

其中算子$\prod\limits_{k=1}^{n}\Delta_{a_kb_k}$定义如下：

$$\Delta_{a_kb_k}G(x)=G(x_1,\cdots,x_{k-1},b_k,x_{k+1},\cdots,x_n)$$
$$-G(x_1,\cdots,x_{k-1},a_k,x_{k+1},\cdots,x_n).$$

$$\prod_{k=1}^{n}\Delta_{a_kb_k}=\Delta_{a_1b_1}\Delta_{a_2b_2}\cdots\Delta_{a_nb_n}.\ (\text{连续作用算子 }\Delta_{a_kb_k}).$$

(条件(iii)来自于当 P 已知时，有

$$\prod_{k=1}^{n}\Delta_{a_kb_k}F(x)=P((a,b])\geqslant 0.$$

定义 1.4.2 若定义在 R^n 上的函数 F 具有性质(i)～(iii)，则称 F 为 R^n 上的分布函数.

与一维情形类似的逆问题是：由分布函数 F 能否唯一决定 $\mathscr{B}^n$ 上的一个概率测度 P，使得(1.4-2)式成立.

定理 1.4.2 设 F 为 R^n 上的一个分布函数，则$\exists$ $(\mathrm{R}^n,\mathscr{B}^n)$上的唯一概率测度 P，$\ni$

$$P((a,b])=\prod_{k=1}^{n}\Delta_{a_kb_k}F(x)\ (\forall\ a,b\in\mathrm{R}^n,\ a\leqslant b).$$

证 类似于一维情形，定义

$$\mathscr{A}_0^n=\{(a,b]:\ a,b\in\mathrm{R}^n,\ a\leqslant b\},$$

则 $\mathscr{A}_0^n$ 为半环，且 R^n 具有 $\mathscr{A}_0^n$-集覆盖. 定义 $\mathscr{A}_0^n$ 上的集函数 P_0：

$$P_0(A)=\prod_{k=1}^{n}\Delta_{a_kb_k}F(x)\ (\forall\ A=(a,b]\in\mathscr{A}_0^n).$$

若 P_0 为 $\mathscr{A}_0^n$ 上的测度(有限的!)，则按推论 1.3.2，P_0 到 $\mathscr{B}^n$ 上的延拓 P 存在且唯一. 而“P_0 为 $\mathscr{A}_0^n$ 上的测度”等价于“P_0 在 $\mathscr{A}_0^n$ 上可加，且上半连续”. 上半连续性之证完全类似于一维情形. 下证

P_0 是可加的.

设 $A=(a^1,a^2]$, $B=(b',b^2]$, $A\bigcup B=(c^1,c^2]$, $A\bigcap B=\varnothing$,则 A 和 B 必有一个$(n-1)$一维的侧面重合.不妨设:

$$\mathop{\times}_{i=2}^{n}(a_i^1,\ a_i^2] = \mathop{\times}_{i=2}^{n}(b_i^1,\ b_i^2],\quad c_1^1=a_1^1,\ c_1^2=b_1^2,\ a_1^2=b_1^1.$$

则

$$\begin{aligned}P_0(A\bigcup B) &= \prod_{k=2}^{n}\Delta_{a_k^1a_k^2}(F(b_1^2,\ x_2,\cdots,x_n)-F(a_1^1,x_2,\cdots,x_n))\\ &=\Delta_{b_1^1b_1^2}\prod_{k=2}^{n}\Delta_{a_k^1a_k^2}F(x_1,x_2,\cdots,x_n)\\ &\quad+\Delta_{a_1^1b_1^1}\prod_{k=2}^{n}\Delta_{a_k^1a_k^2}F(x_1,x_2,\cdots,x_n)\\ &=P_0(A)+P_0(B).\end{aligned}$$

推论 1.4.2 设 F 为 R^n 上的一个分布函数,则$\exists$ $(\mathrm{R}^n,\ \mathscr{B}^n)$上的唯一概率测度 P,使得

$$P((-\infty,x])=F(x)\quad(\forall\ x\in\mathrm{R}^n).$$

证 令 $\mathscr{P}^n=\{(-\infty,x]:\ x\in\mathrm{R}^n\}$,则 $\mathscr{P}^n$ 为 π-类,且 R^n 具有 $\mathscr{P}^n$-集覆盖.定义 $\mathscr{P}^n$ 上的集函数 $\widetilde{P}_0$:

$$\widetilde{P}_0((-\infty,x])=F(x)\quad(\forall\ x\in\mathrm{R}^n),$$

则易证:$\widetilde{P}_0$ 为 $\mathscr{P}^n$ 上的(有限)测度.按定理 1.4.2,$\widetilde{P}_0$ 到 $\mathscr{B}^n=\sigma(\mathscr{P}^n)$上的延拓 P 存在.按定理 1.3.5,此延拓是唯一的.

(三) 可测空间$(\mathrm{R}^\infty,\ \mathscr{B}^\infty)$上的概率测度

设 P 为$(\mathrm{R}^\infty,\ \mathscr{B}^\infty)$上的概率测度,记

$$\mathscr{I}^\infty(B^n)=\{x\in\mathrm{R}^\infty:\ (x_1,x_2,\cdots,x_n)\in B^n\}\quad(B^n\in\mathscr{B}^n).$$

此符号表达 R^∞中具有基 B^n 的柱集.定义 $\mathscr{B}^n$ 上的集函数:

$$P_n(B_n)=P(\mathscr{I}^\infty(B^n))\quad(\forall\ B^n\in\mathscr{B}^n).$$

则 P_n 为$(\mathrm{R}^n,\ \mathscr{B}^n)$上的概率测度,且具有如下性质:

(c)$'$ $\quad P_{n+1}(B^n\times\mathrm{R})=P_n(B^n)\quad(\forall\ n\geqslant1,\ B^n\in\mathscr{B}^n).$

设 $F_n=F_n(x)=P(\mathscr{I}^\infty((-\infty,x]))$ $(\forall\ n\geqslant1,\ x\in\mathrm{R}^n)$，则

(c)″ $F_{n+1}(x_1,x_2,\cdots,x_n,\infty)=F_n(x_1,x_2,\cdots,x_n)$ $(\forall\ n\geqslant1,$ $(x_1,\cdots,x_n)\in\mathrm{R}^n)$.

定义 1.4.3 如果概率测度族 $P=(P_n,\ n\geqslant1)$ 和分布函数族 $F=(F_n,\ n\geqslant1)$ 分别满足条件(c)′和(c)″，则说它们是**相容的**.

定理 1.4.3 设概率测度族 $P=(P_n,\ n\geqslant1)$ 是相容的，则 $\exists$ $(\mathrm{R}^\infty,\ \mathscr{B}^\infty)$ 上的唯一概率测度 P，使得

$$P(\mathscr{I}^\infty(B^n)) = P_n(B^n) \quad (\forall\ n\geqslant1,\ B^n\in\mathscr{B}^n).$$

证 令 $\mathscr{A}^\infty=\{A:\ A=\mathscr{I}^\infty(B^n),\ B^n\in\mathscr{B}^n,\ n\geqslant1\}$，则易证 $\mathscr{A}^\infty$ 为 R^∞ 上的代数. 定义 $\mathscr{A}^\infty$ 上的集函数 P_0:

$$P_0(A) = P_0(\mathscr{I}^\infty(B^n)) = P_n(B^n)$$
$$(\forall\ n\geqslant1,\ A=\mathscr{I}^\infty(B^n)\in\mathscr{A}^\infty).$$

按此式定义的 P_0 在 $\mathscr{A}^\infty$ 上无二意性，即对 $\forall\ A\in\mathscr{A}^\infty$，$P_0(A)$ 之值与 A 的表示无关. 事实上，假如 $\mathscr{I}^\infty(B^n)=A=\mathscr{I}^\infty(C^{n+k})$(某 $k\geqslant0$)，则 $C^{n+k}=B^n\times\mathrm{R}^k$. 于是，按 P 的相容性，有

$$\begin{aligned}P_0(\mathscr{I}^\infty(B^n)) &= P_0(\mathscr{I}^\infty(B^n\times\mathrm{R}^k)) = P_{n+k}(B^n\times\mathrm{R}^k)\\ &= P_{n+k}(C^{n+k}) = P_0(\mathscr{I}^\infty(C^{n+k})).\end{aligned}$$

由此可见，P_0 在 $\mathscr{A}^\infty$ 上的定义无二意性.

假若 P_0 为 $\mathscr{A}^\infty$ 上的有限测度，则按推论 1.3.2，P_0 从 $\mathscr{A}^\infty$ 到 $\mathscr{B}^\infty=\sigma(\mathscr{A}^\infty)$ 上的延拓 P 存在，且唯一，即定理真. 按定理 1.3.3，P_0 为 $\mathscr{A}^\infty$ 上的测度等价于：P_0 在 $\mathscr{A}^\infty$ 上可加，且上半连续. 首先证明：P_0 可加，设 $A_1,A_2\in\mathscr{A}^\infty$，且 $A_1\cap A_2=\varnothing$，则 $A_1=\mathscr{I}^\infty(B_1^n)$，$A_2=\mathscr{I}^\infty(B_2^m)$. 不妨设 $n=m$，由 $A_1A_2=\varnothing$ 知：$B_1^n\cdot B_2^n=\varnothing$. 于是，

$$\begin{aligned}P_0(A_1\cup A_2) &= P_0(\mathscr{I}^\infty(B_1^n)\cup\mathscr{I}^\infty(B_2^n)) = P_0(\mathscr{I}^\infty(B_1^n\cup B_2^n))\\ &= P_n(B_1^n\cup B_2^n) = P_n(B_1^n)+P_n(B_2^n)\\ &= P_0(A_1)+P_0(A_2).\end{aligned}$$

剩下的问题是证明：P_0 上半连续，即

$$(A_n,\ n\geqslant1)\subset\mathscr{A}^\infty,\ 且\ A_n\downarrow\varnothing\Rightarrow P_0(A_n)\downarrow0.$$

采用反证法，假如，$\lim\limits_{n\to\infty}P_0(A_n)=\delta>0$. 则对$\forall\ n\geqslant 1$，存在 $D^n\in\mathscr{B}^n$，

$\overline{于}$ $[D^n]\subset B^n$，且

$$P_n(B^n\backslash D^n)\leqslant\delta\cdot 2^{-(n+1)} \tag{1.4-3}$$

令 $\hat{D}^n=\mathscr{I}^\infty(D^n)$，则由(1.4-3)式知：

$$P_0(A_n\backslash\hat{D}^n)=P_0(\mathscr{I}^\infty(B^n\backslash D^n))=P_n(B^n\backslash D^n)\leqslant\delta\cdot 2^{-(n+1)}. \tag{1.4-4}$$

令 $\hat{C}_n=\bigcap\limits_{i=1}^n\hat{D}^i$，则

$$[\hat{C}_n]=\bigcap_{i=1}^n[\hat{D}^i],\quad [\hat{D}^i]=\mathscr{I}^\infty([D^i]),\quad \hat{C}_n\in\mathscr{A}^\infty,$$

$$[\hat{C}_n]\subset A_n,\quad [\hat{C}_n]\downarrow\varnothing,\quad \hat{C}_n\downarrow\varnothing.$$

按(1.4-4)式，有

$$\begin{aligned}\delta\leqslant P_0(A_n)&=P_0(A_n\backslash\hat{C}_n)+P_0(\hat{C}_n)\\&=P_0(\hat{C}_n)+P_0\Big(\bigcup_{k=1}^n(A_n\backslash\hat{D}^k)\Big)\\&\leqslant P_0(\hat{C}_n)+P_0\Big(\bigcup_{k=1}^n(A_n\backslash\hat{D}^k)\Big)\\&\leqslant P_0(\hat{C}_n)+\sum_{k=1}^nP_0(A_n\backslash\hat{D}^k)\\&\leqslant P_0(\hat{C}_n)+\frac{1}{2}\delta.\end{aligned}$$

$$\Rightarrow P_0(\hat{C}_n)\geqslant\frac{1}{2}\delta>0\Rightarrow[\hat{C}_n]\neq\varnothing\ (\forall\ n\geqslant 1).$$

$$\Rightarrow\bigcap_{n=1}^\infty[\hat{C}_n]\neq\varnothing.$$ 这导致矛盾，故 $\delta=0$.

注 1 上述定理对一般无穷维可测空间可以不成立. 实际上，这个定理的成立与可测空间的拓扑性质有关. 在 $\mathbb{R}^n$ 中子集的紧性是按常规收敛意义定义的. 这是定理 1.4.1～1.4.3 成立的共同点. 为了说明这个性质的重要性，我们来分析下例.

设 $\Omega=(0,1]$，按如下方式引入拓扑. 定义

$$\varphi_n(\omega) = \begin{cases} 1, & \omega \in \left(0, \dfrac{1}{n}\right), \\ 0, & \omega \in \left[\dfrac{1}{n}, 1\right]. \end{cases}$$

$$\mathscr{G}_n = \{A \subset \Omega: A = \{\omega: \varphi_n(\omega) \in B\} (\forall B \in \mathscr{B})\}$$

$$\mathscr{F}_n = \sigma(\mathscr{G}_1 \cup \cdots \cup \mathscr{G}_n), \quad \mathscr{F} = \sigma(\bigcup_n \mathscr{F}_n).$$

则$(\Omega, \mathscr{F})$, $(\Omega, \mathscr{F}_n)$, $n \geqslant 1$,均为可测空间. 在$(\Omega, \mathscr{F}_n)$上定义概率测度

$$P_n(\omega: (\varphi_1(\omega), \varphi_2(\omega), \cdots, \varphi_n(\omega)) \in B^n)$$

$$= \begin{cases} 1, & (1,1,\cdots,1) \in B^n, \\ 0, & \text{其它}. \end{cases} \quad (\forall B^n \in \mathscr{B}^n).$$

此概率测度族$\mathscr{P} = (P_n, n \geqslant 1)$具有下列特性:

(i) $\mathscr{P}$ 是相容的,即 $P_{n+1}|_{\mathscr{F}_n} = P_n$.

(ii) 不存在$(\Omega, \mathscr{F})$上的概率测度 P,使得

$$P|_{\mathscr{F}_n} = P_n \quad (\forall n \geqslant 1).$$

事实上,若$\exists$ P 于$(\Omega, \mathscr{F})$上,$\ni$ $P|_{\mathscr{F}_n} = P_n (\forall n \geqslant 1)$. 则

$$P(\omega: \varphi_i(\omega) = 1, 1 \leqslant i \leqslant n)$$

$$= P_n(\omega: \varphi_i(\omega) = 1, 1 \leqslant i \leqslant n) = 1 \quad (\forall n \geqslant 1).$$

$$\Rightarrow P(\omega: \varphi_i(\omega) = 1, 1 \leqslant i \leqslant n) \xrightarrow[n \to \infty]{} 1.$$

另一方面,令 $A_n = \left(0, \dfrac{1}{n}\right) \downarrow \varnothing$,则按测度的连续性,有 $P(A_n) \downarrow 0$. 然而,$A_n = \bigcap_{k=1}^{n} A_k$,且

$$P(A_n) = P(\bigcap_{k=1}^{n} A_k) = P(\omega: \varphi_k(\omega) = 1, 1 \leqslant k \leqslant n) \xrightarrow[n \to \infty]{} 1.$$

这就产生出矛盾的结果.

现在来考察一下问题出在哪里. 显然,

$$\mathscr{F}_1 = \{\varnothing, \Omega, A_1, A_2\}, \quad A_1 = \{1\}, A_2 = (0,1).$$

在 A_2 中按常规定义收敛的任意紧子集都不是 $\mathscr{F}_1$ 中的元素. 也就是说,对$\forall$ $\delta > 0$,在 A_2 中不存在非空子集 B, $\ni$ $[B] \subset A_2$, $B \in$

$\mathscr{F}_1$, $P(A_2\backslash B)\leqslant\delta$.

推论 1.4.3 设分布函数族 $F=(F_n, n\geqslant 1)$ 满足相容性条件 (c)″. 则存在唯一概率测度 P 于 $(\mathrm{R}^\infty, \mathscr{B}^\infty)$ 上，

$$
\ni \quad P(\mathscr{I}^\infty((-\infty, x^n])) = F_n(x^n) \quad (\forall\, n \geqslant 1,\ x^n \in \mathrm{R}^n) \tag{1.4-5}
$$

证 令 $\mathscr{G}^\infty=\{\mathscr{I}^\infty((-\infty, x^n]): x^n\in\mathrm{R}^n, n\geqslant 1\}$，则 $\mathscr{I}^\infty$ 为 π-类，且 R^∞ 具有 $\mathscr{G}^\infty$-集覆盖. 设 $\widetilde{P}_0$ 由(1.4-5)式定义. 若 $\widetilde{P}_0$ 为 $\mathscr{G}^\infty$ 上的测度. 则按定理 1.3.5，此推论成立. 按定理 1.4.3 之证，定义在 $\mathscr{A}^\infty$ 上的 P_0 可唯一延拓成 $\mathscr{B}^\infty$ 上的测度，且 P_0 为 $\mathscr{A}^\infty$ 上的测度. 由 $\widetilde{P}_0=P_0|_{\mathscr{G}^\infty}$ 知：$\widetilde{P}_0$ 为 $\mathscr{G}^\infty$ 上的概率测度.

（四）可测空间$(\mathrm{R}^T, \mathscr{B}^T)$上的概率测度

设 T 为一有序指标集. R_t 表示相应于 $t\in T$ 的实数集合. $\tau=[t_1,\cdots,t_n]\subset T$ 表示具有 n 个元素的有限无序子集. 记 $|\tau|=n$. $\mathrm{R}^\tau=\mathop{\times}\limits_{k=1}^{|\tau|}\mathrm{R}_{t_k}$. P_τ 为$(\mathrm{R}^\tau, \mathscr{B}^\tau)$上的概率测度.

$$
\mathscr{I}^T(B^\tau) = \{x \in \mathrm{R}^T: (x_{t_1},\cdots,x_{t_n}) = x^\tau \in B^\tau\}. \quad (B^\tau \in \mathscr{B}^\tau).
$$

假定 P 为$(\mathrm{R}^T, \mathscr{B}^T)$上的概率测度. 令

$$
P_\tau(B^\tau) = P(\mathscr{I}^T(B^\tau)) \quad (\forall\, B^\tau \in \mathscr{B}^\tau).
$$

则 $\mathscr{P}=\{P_\tau: \tau\subset T$ 为有限无序子集$\}$为概率测度族.

令 $F_\tau(x^\tau)=P_\tau((-\infty, x^\tau])\ (\forall\, x^\tau\in\mathrm{R}^\tau)$，

则 $F=\{F_\tau: \tau\subset T$ 为有限无序子集$\}$为分布函数族. $\mathscr{P}$，F 具有下列性质：

(c)$_1$ $P_\sigma(x^\sigma: x^\tau\in B^\tau)=P_\tau(B^\tau)$，其中 $\tau,\sigma\subset T$ 为有限无序子集，且 $\tau\subset\sigma$.

(c)$_2$ $F_\sigma(x^\sigma)=F_\tau(x^\tau)$，其中 τ,σ 同(c)$_1$；$x^\tau\in\mathrm{R}^\tau$；$x_t=\infty$ 当 $t\in\sigma\backslash\tau$ 时.

定义 1.4.4 如果概率测度族 $\mathscr{P}$，分布函数族 F 分别满足(c)$_1$ 和(c)$_2$，则称 $\mathscr{P}$，F 是相容的.

以下讨论逆问题.

定理 1.4.4(Kolmogorov 定理). 设概率测度 $\mathscr{P}$ 是相容的，则∃ 唯一概率测度 P 于$(\mathrm{R}^T, \mathscr{B}^T)$上,使得

$$P(\mathscr{I}^T(B^\tau)) = P_\tau(B^\tau) \quad (\forall\ B^\tau \in \mathscr{B}^\tau), \tag{1.4-6}$$

其中 $\tau \subset T$ 为任意有限无序子集.

证 按定理 1.2.1,$\hat{B} \in \mathscr{B}^T$ 具有如下表示:

$$\hat{B} = \mathscr{I}_S^T(B), \quad S = (s_k, k \geqslant 1) \subset T, \quad B \subset \mathscr{B}^S.$$

定义 $\mathscr{B}^T$ 上的集函数

$$P(\hat{B}) \triangleq P(\mathscr{I}_S^T(B)) = P_S(B),$$
$$S = (s_k, k \geqslant 1) \subset T, B \in \mathscr{B}^S \tag{1.4-7}$$

其中 P_S 为$(\mathrm{R}^S, \mathscr{B}^S)$上的概率测度. 按定理 1.4.3,$P_S$ 唯一地由(1.4-6)式决定(当 $\tau \subset S$ 时). 变动 $S \subset T$ 即知:由(1.4-7)式定义的 P 满足条件(1.4-6)式. 剩下的问题就是证明:P 为所要求的概率测度.

首先证明:P 是相容的,即 $P(\hat{B})$的值与 $\hat{B}$ 的表示无关. 事实上,设 $\mathscr{I}_S^T(B^S) = \hat{B} = \mathscr{I}_{S'}^T(B^{S'})$,其中 $S \subset S' \subset T$,则 $B^{S'} = \{x^{S'}: x^S \in B^S\}$. 按相容性条件$(\mathrm{c})_1$,有

$$P(\mathscr{I}_{S'}^T(B^{S'})) = P_{S'}(B^{S'}) = P_S(B^S) = P(\mathscr{I}_S^T(B^S)).$$

现在证明:P 具有可数可加性. 设$(\hat{B}_n, n \geqslant 1) \subset \mathscr{B}^T$ 是不相联的. 则按定理 1.2.1,∃ 可数集 $S \subset T$,及 $B_n^S \in \mathscr{B}^S (n \geqslant 1)$, ∋ $\hat{B}_n = \mathscr{I}_S^T(B_n^S)$ $(\forall\ n \geqslant 1)$. 显然,$(B_n^S, n \geqslant 1)$是不相联的,且 $\bigcup_n \hat{B}_n = \mathscr{I}_S^T(\bigcup_n B_n^S)$. 于是,

$$P(\bigcup_n \hat{B}_n) = P(\mathscr{I}_S^T(\bigcup_n B_n^S)) = P_S(\bigcup_n B_n^S) = \sum_n P_S(B_n^S)$$
$$= \sum_n P(\mathscr{I}_S^T(B_n^S)) = \sum_n P(\hat{B}_n).$$

推论 1.4.4 设分布函数族 F 满足相容性条件$(\mathrm{c})_2$,则存在唯一概率测度 P 于$(\mathrm{R}^T, \mathscr{B}^T)$上,使得

$$P(\mathscr{I}_\tau^T((-\infty, x^\tau])) = F_\tau(x^\tau) \quad (\forall\, x^\tau \in \mathrm{R}^\tau), \quad (1.4\text{-}8)$$

其中 $\tau \subset T$ 为任意有限无序子集.

证 按推论 1.4.2,由 F_τ 可唯一地产生$(\mathrm{R}^\tau, \mathscr{B}^\tau)$上的概率测度 P_τ,使得 $P_\tau((-\infty, x^\tau]) = F_\tau(x^\tau)$. 不难由 F 的相容性推出 $\mathscr{P} = (P_\tau)$的相容性. 按定理 1.4.4,由 $\mathscr{P}$ 可唯一地产生$(\mathrm{R}^T, \mathscr{B}^T)$上的概率测度 P,使得(1.4.8)式成立.

注 2 在此定理中,实线 R_t 可由任意完备可分尺度空间 Ω_t 代替. 相应的 σ-代数 $\mathscr{F}_t$ 可由开集类 $\mathscr{O}_t$ 产生. 假如 P_t 是$(\Omega_t, \mathscr{F}_t)$上的概率测度,则按此拓扑,概率空间$(\Omega_t, \mathscr{F}_t, P_t)$具有如下性质:若 $A \in \mathscr{O}_t$,且 $P_t(A) > 0$,则对$\forall\ \delta > 0$,$\exists$ 开集 $B \subset A$,$\ni$ 闭包$[B] \subset A$,且 $P_t(A \setminus [B]) \leqslant \delta$. 注意,将$\varnothing$也当作开集,即$\varnothing \in \mathscr{O}_t$,则 $\mathscr{O}_t$ 为 π-类,且 $\Omega_t \in \mathscr{O}_t$. 因此,当以$\mathop{\times}\limits_{t \in T}(\Omega_t, \mathscr{F}_t)$代替$(\mathrm{R}^T, \mathscr{B}^T)$时,本节结论之证的各个环节均可通过.

例 1.4.1 设 $T = [0, \infty)$. 定义

$$\varphi_t(y|x) = \begin{cases} \dfrac{1}{\sqrt{2\pi t}} \exp\left(-\dfrac{1}{2t}(y-x)^2\right), & t > 0, \\ \delta(y-x), & t = 0. \end{cases} \quad (x, y \in \mathrm{R}).$$

其中 $\delta(x)$为 R 上的广义函数. 记

$$B = \mathop{\times}_{k=1}^{n} I_k,\ I_k = (a_k, b_k] \quad (a_k, b_k \in \mathrm{R},\ a_k \leqslant b_k)$$

$$\tau = [t_1, \cdots, t_n] \subset T,\ 0 \leqslant t_1 < t_2 < \cdots < t_n.$$

则按定理 1.4.2,存在唯一概率测度 P_τ 于$(\mathrm{R}^\tau, \mathscr{B}^\tau)$上,$\ni$

$$P_\tau(B) = \int_{I_n} \cdots \int_{I_1} \varphi_{t_1-0}(y_1|0)\varphi_{t_2-t_1}(y_2|y_1) \cdots \varphi_{t_n-t_{n-1}}(y_n|y_{n-1})\mathrm{d}y_1 \cdots \mathrm{d}y_n.$$

不难验证,$\mathscr{P} = (P_\tau)$满足相容性条件$(\mathrm{c})_1$. 按定理 1.4.4,$\mathscr{P}$ 唯一地产生$(\mathrm{R}^T, \mathscr{B}^T)$上的一个概率测度 P,$\ni$ P 具有族 $\mathscr{P}$. 通常称此测度 P 为 Wiener 测度.

§ 1.5 Kolmogorov 过程构造定理

首先介绍随机元概念，它概括任意类型的随机对象.

定义 1.5.1 设$(\Omega, \mathscr{F})$, $(E, \mathscr{E})$为可测空间，ξ是$\Omega \to E$的映射. 若

$$\xi^{-1}(B) \triangleq \{\omega: \xi(\omega) \in B\} \in \mathscr{F} \quad (\forall B \in \mathscr{E}),$$

则称ξ是$\mathscr{F} \backslash \mathscr{E}$-**可测的**. 这类可测映射统称为**随机元**. 有时也称为E-值 r.v. (r.v. =Random variable).

在随机元中有若干种很重要的特别对象：

(1) $(E, \mathscr{E}) = (\mathrm{R}, \mathscr{B})$ (称ξ为 r.v.)；

(2) $(E, \mathscr{E}) = (\mathrm{R}^n, \mathscr{B}^n)$ (称ξ为随机矢量或随机点)；

(3) $(E, \mathscr{E}) = (\mathrm{R}^\infty, \mathscr{B}^\infty)$ (称ξ为随机序列)；

(4) $(E, \mathscr{E}) = (\mathrm{R}^T, \mathscr{B}^T)$ (称ξ为随机过程).

映射ξ: $(\mathrm{R}, \mathscr{B}) \to (\mathrm{R}, \mathscr{B})$一般称为 Borel 函数. 而$n$元实值 Borel 函数$\xi$就是指可测映射$\xi$:

$$(\mathrm{R}^n, \mathscr{B}^n) \to (\mathrm{R}, \mathscr{B}).$$

引理 1.5.1 设$\mathscr{A}$为R上的某子集类，且$\mathscr{B} = \sigma(\mathscr{A})$. ξ是$\Omega \to \mathrm{R}$的映射，则ξ为 r.v. 的充要条件是：

$$\xi^{-1}(A) \in \mathscr{F} \quad (\forall A \in \mathscr{A}).$$

证 必要性显然. 下证充分性. 令

$$\mathscr{G} = \{B: \xi^{-1}(B) \in \mathscr{F}, B \in \mathscr{B}\}.$$

则$\mathscr{A} \subset \mathscr{G} \subset \mathscr{B}$. 因此，若$\mathscr{G}$为$\sigma$-代数，则必有$\mathscr{G} \subset \mathscr{B}$. 这表明：$\xi$为 r.v. 下面验证$\mathscr{G}$为$\sigma$-代数.

设$B \in \mathscr{G}$，则$\xi^{-1}(B) \in \mathscr{F}$. 注意，$\xi^{-1}(\mathrm{R}) = \Omega \in \mathscr{F}$. 因此，

$$\xi^{-1}(B^c) = \xi^{-1}(\mathrm{R}) \backslash \xi^{-1}(B) = (\xi^{-1}(B))^c \in \mathscr{F}.$$

从而，$B^c \in \mathscr{G}$. 设$B_n \in \mathscr{G}$ $(n \geqslant 1)$，则显然有

$$\xi^{-1}(\bigcup_n B_n) = \bigcup_n \xi^{-1}(B_n) \in \mathscr{F} \Rightarrow \bigcup_n B_n \in \mathscr{G}.$$

故$\mathscr{G}$为σ-代数.

推论 1.5.1 设ξ为$\Omega\to\mathrm{R}$的映射. ξ为r.v.的充要条件是:
$\{\omega: \xi(\omega)\leqslant x\}\in\mathscr{F}$ 或 $\{\omega: \xi(\omega)<x\}\in\mathscr{F}$ $(\forall\ x\in\mathrm{R})$.

证 令$\mathscr{A}_1=\{B: B=(-\infty,x],\ x\in\mathrm{R}\}$,

$\mathscr{A}_2=\{B: B=(-\infty,x),\ x\in\mathrm{R}\}$,

则$\sigma(\mathscr{A}_i)=\mathscr{B}$ $(i=1,2)$. 于是,此推论为引理1.5.1的特例.

引理 1.5.2 设φ为$\mathrm{R}\to\mathrm{R}$的Borel函数. ξ为$(\Omega,\mathscr{F})$上的r.v. 则复合映射$\eta=\varphi\circ\xi$,即$\eta(\omega)=\varphi(\xi(\omega))$,同样是r.v.

证 由$\xi^{-1}(\varphi^{-1}(B))=\eta^{-1}(B)$立即可得此结论.

假定ξ为$(\Omega,\mathscr{F})\to(E,\mathscr{E})$的随机元. 定义

$$\mathscr{F}^{\xi}=\{A: A=\xi^{-1}(B),\ B\in\mathscr{E}\}.$$

则不难验证:$\mathscr{F}^{\xi}$为σ-代数. 每个随机元ξ都可生成一个相应的σ-代数$\mathscr{F}^{\xi}$. 以后,"$\mathscr{F}\backslash\mathscr{E}$-可测"简写为"$\mathscr{F}$-可测". 下列定理统称为可测函数的表现定理.

定理 1.5.1 设ξ为m-维随机矢量. η为r.v. 则η是$\mathscr{F}^{\xi}$-可测的充要条件是:$\exists$ m元Borel函数φ,使得

$$\eta=\varphi\circ\xi.$$

证 充分性显然. 下证必要性. 为此,定义

$\Phi_{\xi}=\{\gamma: \gamma$ 是 $\mathscr{F}^{\xi}$- 可测的$\}$,

$\Phi_{\xi}^{*}=\{\gamma: \gamma=\psi\circ\xi$,其中$\psi$为$m$元Borel函数$\}$.

显然,$\Phi_{\xi}^{*}\subset\Phi_{\xi}$. 于是,必要性等价于:$\Phi_{\xi}\subset\Phi_{\xi}^{*}$.

设$A\in\mathscr{F}^{\xi}$,$\eta(\omega)=1_A(\omega)=\begin{cases}1, & \omega\in A,\\ 0, & \omega\notin A.\end{cases}$ 则$\eta\in\Phi_{\xi}^{*}$. 事实上,按$\mathscr{F}^{\xi}$之定义,$\exists$ $B\in\mathscr{B}^m$, 使 $A=\xi^{-1}(B)$. 令$\varphi(x)=1_B(x)$ $(x\in R^m)$,则φ为m元Borel函数,且$\eta=1_A=\varphi\circ\xi.\Rightarrow\eta\in\Phi_{\xi}^{*}$.

若$\eta=\sum\limits_{i=1}^{n}C_i1_{A_i}$, $\psi=\sum\limits_{i=1}^{n}C_i1_{B_i}$,其中$A_i=\xi^{-1}(B_i)$. 则按前一段之推理,易证:$\eta=\psi\circ\xi\in\Phi_{\xi}^{*}$.

现设η为任意$\mathscr{F}^{\xi}$-可测的r.v.,即$\eta\in\Phi_{\xi}$. 令

$$A_{nk}=\left\{\omega: \frac{k-1}{2^n}\leqslant\eta<\frac{k}{2^n}\right\},\ A_n=\{\omega: n\leqslant\eta\},$$

$$A_{-n} = \{\omega: \eta < -n\},$$

$$\eta_n(\omega) = \sum_{k=-n2^n}^{n2^n} \frac{k-1}{2^n} 1_{A_{nk}}(\omega) + n1_{A_n}(\omega) - n \cdot 1_{A_{-n}}(\omega).$$

则由 η 为 r. v. 知：$\lim\limits_{n\to\infty}\eta_n(\omega)=\eta(\omega)$ （$\forall\ \omega\in\Omega$），且 $\eta_n\in\Phi_\xi^*$ ，于是，对 $\forall\ n\geqslant 1$，$\exists\ m$ 元 Borel 函数 φ_n，$\ni\ \eta_n=\varphi_n\circ\xi$，且

$$\lim_{n\to\infty}\varphi_n\circ\xi(\omega) = \eta(\omega).$$

令 $B=\{x\in \mathrm{R}^m: \lim\limits_{n\to\infty}\varphi_n(x)$存在$\}$，则

$$B^c = \mathrm{R}^m \backslash B = \bigcup_{k=1}^{\infty}\bigcap_{u=1}^{\infty}\bigcup_{l,n=1}^{\infty}\left\{x\in\mathrm{R}^m: |\varphi_{u+l}(x)-\varphi_{u+n}(x)|\geqslant\frac{1}{k}\right\}.$$

从而，$B^c, B\in\mathscr{B}^m$. 令

$$\varphi(x) = \begin{cases}\lim\limits_{n\to\infty}\varphi_n(x), & x\in B,\\ 0, & x\notin B.\end{cases}$$

则 φ 为 m 元 Borel 函数，且

$$\eta(\omega) = \lim_{n\to\infty}\varphi_n(\xi(\omega)) = \varphi(\xi(\omega)) = \varphi\circ\xi(\omega)$$

$$\Rightarrow \eta\in\Phi_\xi^*.$$

推论 1.5.2 设 $\mathscr{D}=(D_n,\ n\geqslant 1)$为 Ω 的一个可数分解，且 $\mathscr{F}=\sigma(\mathscr{D})$. 若 η 是$(\Omega,\mathscr{F})$上的一个 r. v. 则 η 具有表示：$\eta=\sum\limits_{k=1}^{\infty}x_k 1_{D_k}$.

证 设 $\xi=\sum\limits_{k=1}^{\infty}k\cdot 1_{D_k}$，则 ξ 是$(\Omega,\mathscr{F})\to(\mathrm{R},\mathscr{B})$可测的，即 ξ 为 r. v.，且 $\mathscr{F}^\xi=\mathscr{F}$. 于是，按定理 1.5.1，$\exists$ Borel 函数 φ，$\ni\ \eta=\varphi\circ\xi$，而 $\varphi\circ\xi$ 有表现形式：

$$\varphi\circ\xi(\omega) = \varphi(k)\ \text{当}\ \omega\in D_k\ \text{时}.$$

令 $x_k=\varphi(k)$便得 η 的表示.

定理 1.5.2 设 ξ 是$(\Omega,\mathscr{F})\to(\mathrm{R}^T,\mathscr{B}^T)$的随机元，其中 T 为有序参数集. η 为$(\Omega,\mathscr{F})$上的 r. v. 则 η 是 $\mathscr{F}^\xi$-可测的充要条件是：

存在可测映射 $f:(\mathrm{R}^\infty,\mathscr{B}^\infty)\to(\mathrm{R},\mathscr{B})$ 及可数集$(t_k,\ k\geqslant 1)$，使

$$\eta(\omega)=f(\xi_{t_1}(\omega),\ \xi_{t_2}(\omega),\ \cdots)\ (\forall\ \omega\in\Omega). \qquad (1.5\text{-}1)$$

证 充分性显然，下证必要性.

首先设 $\eta=1_A$，$A\in\mathscr{F}^\xi$. 这时，$\exists\ B\in\mathscr{B}^T$，使得 $A=\xi^{-1}(B)$. 按定理 1.2.1，$\exists\ (t_k,\ k\geqslant 1)\subset T$ 及 $B^*\in\mathscr{B}^\infty$，使得

$$B=\{x\in\mathrm{R}^T:\ (x_{t_1},\ x_{t_2},\ \cdots)\in B^*\}.$$

令 $\sigma=(t_k,\ k\geqslant 1)$，$x^\sigma=(x_{t_1},\ x_{t_2},\ \cdots)$，则

$$\eta(\omega)=1_B(\xi(\omega))=1_{B^*}(\xi^\sigma(\omega)).$$

显然，$f(x)=1_{B^*}(x)\ (x\in\mathrm{R}^\infty)$是可测的. 故(1.5-1)式成立. 由此不难证明，对 η 为简单映射的情形，(1.5-1)式成立.

现设 η 非负. 令 $A_n=\{\omega:\ n\leqslant\eta\}$，$A_{nk}=\left\{\omega:\ \frac{k-1}{2^n}\leqslant\eta<\frac{k}{2^n}\right\}$，

$$\eta_n=\sum_{k=1}^{n2^n}\frac{k-1}{2^n}1_{A_{nk}}+n1_{A_n}.$$

则 $A_{nk},A_n\in\mathscr{F}^\xi$，$\eta_n$ 是 $\mathscr{F}^\xi$-可测的. 且 η_n 具有表现(1.5-1)式及 $\eta(\omega)=\lim\limits_{n\to\infty}\eta_n(\omega)\ (\forall\ \omega\in\Omega)$. 按前面的证明：$\exists\ \sigma_{nk}=(t_i^{nk},\ i\geqslant 1)\subset T$，$\sigma_n=(t_i^n,\ i\geqslant 1)\subset T$ 及可测映射 $\varphi_{nk},\varphi_n:\ (\mathrm{R}^\infty,\ \mathscr{B}^\infty)\to(\mathrm{R},\ \mathscr{B})$，使

$$\eta_n(\omega)=\sum_{k=1}^{n2^n}\varphi_{nk}(\xi^{\sigma_{nk}}(\omega))+\varphi_n(\xi^{\sigma_n}(\omega)).$$

令 $\sigma=\bigcup\limits_{n=1}^{\infty}\bigcup\limits_{k=1}^{n2^n}(\sigma_n\cup\sigma_{nk})$，并定义投影算子 π_{nk}，π_n：$x^{\sigma_{nk}}=\pi_{nk}(x^\sigma)$，$x^{\sigma_n}=\pi_n(x^\sigma)$，则 $\eta_n(\omega)=f_n\circ\xi^\sigma(\omega)$. 这里 $f_n=\sum\limits_{k=1}^{n2^n}\varphi_{nk}\circ\pi_{nk}+\varphi_n\circ\pi_n$.

令

$$B=\{x\in\mathrm{R}^\infty:\ \lim_{n\to\infty}f_n(x)\ 存在\},$$

则易证：$B\in\mathscr{B}^\infty$. 定义映射 f：

$$f(x)=\begin{cases}\lim\limits_{n\to\infty}f_n(x), & x\in B,\\ 0, & x\notin B.\end{cases}\quad(\forall\ x\in\mathrm{R}^\infty).$$

则 f 是 $\mathscr{B}^\infty$-可测的，且

$$\eta(\omega) = \lim_{n\to\infty} f_n \circ \xi^\sigma(\omega) = f \circ \xi^\sigma(\omega) \quad (\forall\ \omega \in \Omega).$$

这表明：当 η 非负时，(1. 5-1)式成立.

最后考虑一般的 r. v. η，这时可有分解：

$$\eta = \eta^+ - \eta^-, \quad \eta^+ = \eta \vee 0, \quad \eta^- = -\eta \wedge 0.$$

则 $\eta^\pm$ 为非负 r. v. 且 $\mathscr{F}^\xi$-可测. 按上面之证，

存在可数集 $\sigma_1, \sigma_2 \subset T$ 及 $\mathscr{B}^\infty$-可测的函数 f_1, f_2,

使得 $\eta^+ = f_1 \circ \xi^{\sigma_1}$, $\quad \eta^- = f_2 \circ \xi^{\sigma_2}$.

令 $\sigma = \sigma_1 \cup \sigma_2$，则 σ 为 T 中的一个可数集. 定义投影算子 π_1, π_2： $\pi_i(x^\sigma) = x^{\sigma_i}$ $(i=1,2)$，并令 $f^+ = f_1 \circ \pi_1$, $f^- = f_2 \circ \pi_2$，则 $\eta = f \circ \xi^\sigma$. 即(1. 5-1)式成立.

下面介绍 Kolmogorov 过程构造定理，它在过程论中起着特别重要的指导作用.

定义 1. 5. 2 设 $(\Omega, \mathscr{F}, P)$ 为某概率空间，$\xi = (\xi_t,\ t \in T)$ 为 $(\Omega, \mathscr{F}) \to (\mathrm{R}^T,\ \mathscr{B}^T)$ 的随机过程. 定义

$$P_\xi(B) = P(\xi^{-1}(B)) \quad (\forall\ B \in \mathscr{B}^T),$$

则称 P_ξ 是由 ξ 诱导的 $(\mathrm{R}^T,\ \mathscr{B}^T)$ 上的概率分布. 设 τ 为 T 中任意有限无序子集. 定义

$$P_\tau(B) = P(\omega:\ \xi^\tau(\omega) \in B) \quad (\forall\ B \in \mathscr{B}^\tau).$$

则称 $P_\xi = (P_\tau,\ \tau \subset T)$ 为 ξ 的有限维概率分布族. 定义函数 $F_\tau(x) = P(\omega:\ \xi^\tau(\omega) \leqslant x^\tau)$ $(\forall\ x^\tau \in R^\tau)$，则称 $F_\xi = (F_\tau,\ \tau \subset T)$ 为 ξ 的有限维分布函数族.

现在提出一个反问题：假如 $P = (P_\tau : \tau \subset T)$, $F = (F_\tau : \tau \subset T)$ 分别为已知的概率分布族和分布函数族. 试问：能否找到一个概率空间 $(\Omega, \mathscr{F}, P)$，及定义于其上的随机过程 ξ，使得

$$P_\xi = P \quad 或 \quad F_\xi = F.$$

定理 1. 5. 3(K-过程构造定理). 设 $F = (F_\tau, \tau \subset T)$ 满足 § 4 中的相容性条件(c)″，则存在一个概率空间 $(\Omega, \mathscr{F}, P)$ 及定义于其

上的随机过程 $\xi=(\xi_t,\ t\in T)$，使 $F_\xi=F$，其中 $F_\xi=(F^\xi_\tau,\ \tau\subset T)$ 是由 ξ 诱导出的分布函数族，即 $F^\xi_\tau(x)=F_\tau(x)\ (x\in \mathrm{R}^\tau,\ \tau\subset T)$.
（注意：τ 表示 T 中的有限无序子集）.

证 选取 $(\Omega,\mathscr{F})=(\mathrm{R}^T,\ \mathscr{B}^T)$. 按推论 1.4.4，存在唯一概率测度 P 于 $(\mathrm{R}^T,\ \mathscr{B}^T)$ 上，使得

$$P(\mathscr{I}^T_\tau((-\infty,x^\tau]))=F_\tau(x^\tau)\quad(\forall\ x^\tau\in \mathrm{R}^\tau).$$

这就作出了所要的概率空间 $(\Omega,\mathscr{F},P)$. 现在定义过程 $\xi=(\xi_t,\ t\in T)$，其中 $\xi_t(\omega)=\omega_t(\forall\ t\in T,\ \omega\in\Omega)$. 这里 ω_t 表示 ω 相应于 t 的坐标分量，则

$$\begin{aligned}F^\xi_\tau(x^\tau)&=P(\omega;\ \xi^\tau(\omega)\leqslant x^\tau)=P(\mathscr{I}^T_\tau((-\infty,x^\tau]))\\&=F_\tau(x^\tau)\quad(\forall\ x^\tau\in \mathrm{R}^\tau).\end{aligned}$$

故 ξ 即为所求的随机过程.

注 1 若将 F 改成 P，(c)″改成(c)′，则上述定理亦成立. 这时依据的不是推论 1.4.4，而是定理 1.4.4.

注 2 上述定理解决了过程的存在性问题. 显然，问题的解答不唯一. 若过程族 $(\Omega^k,\ \mathscr{F}^k,\ P^k,\ \xi^k)$，$k\in K$（指标集），都具有同一有限维分布族 F 或 P，则称此族中的元素是等价的，其中任何一个都可以作为它们的代表. 以后，涉及唯一性时即按此等价意义理解.

注 3 上述定理成立依据于推论 1.4.4（定理 1.4.4）. 这实际上用到了 R 上的自然拓扑结构. 假如不考虑这一特性，上述定理可以不成立. 因此，与拓扑结构无关的情形只可能出现在概率测度族 P 具有某种特别的结构. 著名的 Ionescu Tulcea 定理就是从这一角度出发所得到的一个重要的结果.

定理 1.5.4（I. T. -定理）. 设 $(\Omega_n,\ \mathscr{F}_n)$，$n\geqslant 1$，为可测空间列. $(\Omega,\ \mathscr{F})=\mathop{\times}\limits_{n=1}^{\infty}(\Omega_n,\ \mathscr{F}_n)$. P_1 为 $(\Omega_1,\ \mathscr{F}_1)$ 上的概率测度. $P(\omega_1,\ \omega_2,\ \cdots,\ \omega_n;\ B)$ 满足下列要求：

(i) 对 $\forall\ (\omega_1,\ \omega_2,\ \cdots,\ \omega_n)\in\mathop{\times}\limits_{k=1}^{n}\Omega_k$，$P(\omega_1,\ \omega_2,\ \cdots,\ \omega_n;\ \cdot\ )$ 为

$\mathscr{F}_{n+1}$上的概率测度.

(ii) 对$\forall\ B\in\mathscr{F}_{n+1}$, $P(\cdot\,;B)$为$\mathop{\times}\limits_{k=1}^{n}(\Omega_k,\mathscr{F}_k)$上的 r. v.

定义如下集函数 P_n:

$$P_n(\mathop{\times}\limits_{k=1}^{n}A_k)=\int_{A_1}P_1(\mathrm{d}\omega_1)\int_{A_2}P(\omega_1;\mathrm{d}\omega_2)$$

$$\cdots\int_{A_n}P(\omega_1,\omega_2,\cdots,\omega_n;\mathrm{d}\omega_n),\qquad(1.5\text{-}2)$$

其中$\mathop{\times}\limits_{k=1}^{n}A_k\in\mathop{\times}\limits_{k=1}^{n}\mathscr{F}_k$.

则$\exists$ 唯一概率测度 P 于$(\Omega,\mathscr{F})$上,使得

$$P(\mathscr{I}^{\infty}(\mathop{\times}\limits_{k=1}^{n}A_k))=P_n(\mathop{\times}\limits_{k=1}^{n}A_k)\quad(\forall\ n\geqslant1,\ \mathop{\times}\limits_{k=1}^{n}A_k\in\mathop{\times}\limits_{k=1}^{n}\mathscr{F}_k).$$

且$\exists$ 一个概率空间$(\widetilde{\Omega},\widetilde{\mathscr{F}},\widetilde{P})$及定义于其上的 Ω-值 r. v. η,使得

$$\widetilde{P}(\widetilde{\omega}:\eta(\widetilde{\omega})\in\mathscr{I}^{\infty}(\mathop{\times}\limits_{k=1}^{n}A_k))=P_n(\mathop{\times}\limits_{k=1}^{n}A_k)$$

$$(\forall\ n\geqslant1,\ \mathop{\times}\limits_{k=1}^{n}A_k\in\mathop{\times}\limits_{k=1}^{n}\mathscr{F}_k).$$

$$\mathscr{I}^{\infty}(A)=\{\omega\in\Omega:(\omega_1,\omega_2,\cdots,\omega_n)\in A\}\quad(A\in\mathop{\times}\limits_{k=1}^{n}\mathscr{F}_k).$$

证 令

$$\mathscr{A}_0^n\doteq\{A:A=\mathop{\times}\limits_{k=1}^{n}A_k,\ A_k\in\mathscr{F}_k\}.$$

$$\mathscr{A}_1^n=\{B:B=\bigcup_{l=1}^{m}B_l,\ B_l\in\mathscr{A}_0^n,\ m\geqslant1\}.$$

则 $\mathscr{A}_0^n$, $\mathscr{A}_1^n$ 分别为$\mathop{\times}\limits_{k=1}^{n}\Omega_k$ 上的半代数和代数,且

$$\mathop{\times}\limits_{k=1}^{n}\mathscr{F}_k=\sigma(\mathscr{A}_1^n)=\sigma(\mathscr{A}_0^n).$$

定义$\mathop{\times}\limits_{1}^{n}\mathscr{F}_k$ 上的集函数 P_n^*:

$$P_n^*(A)=\int_{\mathop{\times}\limits_{1}^{n}\Omega_k}1_A(\omega_1,\omega_2,\cdots,\omega_n)P_1(\mathrm{d}\omega_1)P(\omega_1;\mathrm{d}\omega_2)$$

$$\cdots P(\omega_1,\omega_2,\cdots,\omega_{n-1};\mathrm{d}\omega_n),$$

其中 $A\in\mathop{\times}\limits_1^n\mathscr{F}_k$. 则 P_n^* 为 $\mathop{\times}\limits_1^n\mathscr{F}_k$ 上的概率测度,且

$$P_n^*\mid_{\mathscr{A}_0^n}=P_n.$$

事实上,当 $n=1$ 时,有 $P_1^*=P_1$. 因此,按假设,P_1^* 为概率测度,且满足(1.5-2)式. 当 $n=2$ 时,按条件(i)和(ii),$\int_{\Omega_2}1_A(\omega_1,\omega_2)P(\omega_1;\mathrm{d}\omega_2)$ 是 $\mathscr{F}_1$-可测的. 从而,二重积分 $\int_{\Omega_1}P_1(\mathrm{d}\omega_1)\int_{\Omega_2}1_A(\omega_1,\omega_2)P(\omega_1;\mathrm{d}\omega_2)$ 存在. 按 Fubini 定理,P_2^* 为 $\mathop{\times}\limits_1^2\mathscr{F}_k$ 上的概率测度,且满足(1.5-2)式. 采用归纳法,便可证明:对$\forall\ n\geqslant1$,P_n^* 为 $\mathop{\times}\limits_1^n\mathscr{F}_k$ 上满足(1.5-2)式的概率测度. 记 $P^*=(P_n^*,\ n\geqslant1)$. 则易证:$P^*$ 满足 §4 中的相容性条件(c)′. 按定理 1.3.5,P_n^* 是 $\mathop{\times}\limits_1^n\mathscr{F}_k$ 上满足(1.5-2)式的唯一概率测度.

记 $\mathscr{A}_0^\infty=\{\hat{A}:\hat{A}=\mathscr{I}^\infty(A),\ A\in\mathop{\times}\limits_1^n\mathscr{F}_k,\ n\geqslant1\}$,则 $\mathscr{A}_0^\infty$ 为代数. 事实上,设 $\hat{A}_i\in\mathscr{A}_0^\infty(i=1,2)$,则有 $\hat{A}_i=\mathscr{I}^\infty(A_i)$,其中 $A_i\in\mathop{\times}\limits_1^{n_i}\mathscr{F}_k(i=1,2)$. 不妨设 $n_1<n_2$,则

$$\hat{A}_1\cup\hat{A}_2=\mathscr{I}^\infty(A_2\cup(A_1\times\mathop{\times}\limits_{n_1+1}^{n_2}\Omega_k))\in\mathscr{A}_0^\infty.$$

$\mathscr{A}_0^\infty$ 封闭于补运算是显然的. 故 $\mathscr{A}_0^\infty$ 为代数.

现在定义 $\mathscr{A}_0^\infty$ 上的集函数:

$$P(\hat{A})=P_n^*(A)\quad(\forall\ n\geqslant1,\ A\in\mathop{\times}\limits_1^n\mathscr{F}_k).\qquad(1.5\text{-}3)$$

由 P^* 的相容性知:(1.5-3)式定义的 P 在 $\hat{A}$ 上的值与 $\hat{A}$ 的表示无关. 由 P^* 的相容性及 $\mathscr{A}_0^\infty$ 的结构不难推出 P 在 $\mathscr{A}_0^\infty$ 上可加. 按定理 1.3.3,P 为 $\mathscr{A}_0^\infty$ 上的测度等价于

$$P(\hat{A}_n)\downarrow0,\text{其中 }\hat{A}_n\downarrow\varnothing\ (\hat{A}_n\in\mathscr{A}_0^\infty).\qquad(1.5\text{-}4)$$

若 P 为 $\mathscr{A}_0^\infty$ 上的测度,则按推论 1.3.2,由(1.5-3)式定义的 P 便

可唯一地延拓成$(\Omega, \mathscr{F})$上的测度. 仍记为P. 这时,选取$(\tilde{\Omega}, \tilde{\mathscr{F}}, \tilde{P})=(\Omega, \mathscr{F}, P)$, $\eta_n(\tilde{\omega})=\tilde{\omega}_n(\forall\ n\geqslant 1, \tilde{\omega}\in\tilde{\Omega})$. 则$(\tilde{\Omega}, \tilde{\mathscr{F}}, \tilde{P}, \eta)$为满足定理中要求的$\Omega$-值 r. v. 列.

剩下的问题就是证明(1.5-4)式成立. 不妨设$\hat{A}_n=\mathscr{I}^\infty(A_n)$, $A_n\in\mathop{\times}\limits_1^n\mathscr{F}_k$. 采用反证法,假定$\lim\limits_{n\to\infty}P(\hat{A}_n)=\delta>0$.

令 $$f_n^{(1)}(\omega_1)=\int_{\Omega_2}P(\omega_1;\ d\omega_2)\cdots\int_{\Omega_n}1_{A_n}(\omega_1,\ \cdots,\ \omega_n)P(\omega_1,\ \cdots,\ \omega_{n-1};\ d\omega_n),$$

则 $$P(\hat{A}_n)=P_n^*(A_n)=\int_{\Omega_1}f_n^{(1)}(\omega_1)P_1(d\omega_1).$$

由$\hat{A}_{n+1}\subset\hat{A}_n$知:$A_{n+1}\subset A_n\times\Omega_{n+1}$,因此,有

$$1_{A_{n+1}}(\omega_1,\ \cdots,\ \omega_{n+1})\leqslant 1_{A_n\times\Omega_{n+1}}(\omega_1,\ \cdots,\ \omega_{n+1})=1_{A_n}(\omega_1,\ \cdots,\ \omega_n).$$

由此知: $f_n^{(1)}(\omega_1)$单调不增. 令$f^{(1)}(\omega_1)=\lim\limits_{n\to\infty}f_n^{(1)}(\omega_1)$. 按 Lebesque 控制收敛定理,有

$$0<\delta=\delta_0=\lim_{n\to\infty}P(\hat{A}_n)=\lim_{n\to\infty}\int_{\Omega_1}f_n^{(1)}(\omega_1)P(d\omega_1)$$
$$=\int_{\Omega_1}f^{(1)}(\omega_1)P_1(d\omega_1).$$

从而,$\exists\ \omega_1^0\in A_1$, 使 $f^{(1)}(\omega_1^0)>0$. 事实上,对$\forall\ \omega_1\notin A_1$,因$A_2\subset A_1\times\Omega_2$必有$(\omega_1,\ \omega_2)\notin A_2$. 依此类推,对$\forall\ n\geqslant 1$,有$(\omega_1,\ \cdots,\ \omega_n)\notin A_n$. 于是,按$f_n^{(1)}(\omega_1)$之定义,有$f_n^{(1)}(\omega_1)=0\ (\forall\ n\geqslant 1)$. 从而, $f^{(1)}(\omega_1)=0\ (\forall\ \omega_1\notin A_1)$. 故所要求的$\omega_1^0\in A_1$是存在的. 注意,当$n\geqslant 2$时,有

$$0<f^{(1)}(\omega_1^0)=\delta_1\leqslant f_n^{(1)}(\omega_1^0)=\int_{\Omega_2}f_n^{(2)}(\omega_2)P(\omega_1^0;\ d\omega_2).$$

类似地讨论,可得$f_n^{(2)}(\omega_2)\downarrow f^{(2)}(\omega_2)$,并存在$(\omega_1^0,\ \omega_2^0)\in A_2$,使得 $\delta_2=f^{(2)}(\omega_2^0)>0$. 继续这一作法,便可得一点$\omega^0=(\omega_1^0,\ \omega_2^0,\ \cdots)$具有下列性质: $(\omega_1^0,\ \omega_2^0,\ \cdots,\omega_n^0)\in A_n(\forall\ n\geqslant 1)$; $\omega^0\in\bigcap\limits_1^\infty\mathscr{I}^\infty(A_n)=$

$\bigcap_1^{\infty} \hat{A}_n$. 然而，这与$\bigcap_n \hat{A}_n = \varnothing$矛盾. 故$\delta=0$.

推论 1.5.3 设$p(x,B)$具有下列性质：

(i) 对$\forall\ x\in \mathrm{R}$，$p(x,B)$是$\mathscr{B}$上的概率测度.

(ii) 对$\forall\ B\in\mathscr{B}$，$p(x,B)$是R上的Borel函数.

此外，让π为$(\mathrm{R},\mathscr{B})$上的概率分布，则$\exists$ 一个概率空间$(\Omega,\ \mathscr{F},\ P)$及定义于其上的r.v. $\xi=(\xi_n,\ n\geqslant 1)$，使得

$$
\begin{aligned}
&P(\omega;\ \xi(\omega)\in \mathscr{I}^{\infty}((-\infty,x^n]))\\
&\quad=\int_{-\infty}^{x_1}\pi(\mathrm{d}u_1)\int_{-\infty}^{x_2}p(u_1,\ \mathrm{d}u_2)\cdots\int_{-\infty}^{x_n}p(u_{n-1},\ \mathrm{d}u_n),
\end{aligned}
$$

其中$n\geqslant 1$，$x^n=(x_1,x_2,\cdots,x_n)\in\mathrm{R}^n$，

$$
\mathscr{I}^{\infty}(B)=\{x\in\mathrm{R}^{\infty}:\ (x_1,\ x_2,\ \cdots,\ x_n)\in B\}\quad(B\in\mathscr{B}^n).
$$

证 设$P_1=\pi$，$(\Omega_k,\ \mathscr{F}_k)=(\mathrm{R},\ \mathscr{B})$，$(\Omega,\ \mathscr{F})=(\mathrm{R}^{\infty},\ \mathscr{B}^{\infty})$，

$$
\begin{aligned}
P(\omega_1,\omega_2,\cdots,\omega_n;\ B)&=p(\omega_n,B)\ (B\in\mathscr{F}_{n+1}\\
&=\mathscr{B},\ (\omega_1,\omega_2,\cdots,\omega_n)\in\mathrm{R}^n).
\end{aligned}
$$

则I.-C. 定理中的条件都满足. 故结论成立.

注4 I.-C. 定理完全不考虑空间中的拓扑性质，只用到积分论中的某些结果. Markov链和Markov过程既可依据Kolmogorov定理也可以依据I.-C. 定理展开.

第二章　概率测度的收敛性

收敛性是现代概率论中最基本的问题.人们从不同的目的出发可以引出各种各样的收敛性概念.然而,在概率测度的收敛性问题中,最常用的有四种类型.本章将介绍这四种收敛类型的概念及它们的主要结果.

§2.1　积　　分

定义 2.1.1　设$(\Omega, \mathscr{F}, \mu)$为测度空间. f 为$(\Omega, \mathscr{F})\to(\overline{\mathrm{R}}, \overline{\mathscr{B}})$的可测函数.这里$\overline{\mathrm{R}}=[-\infty, +\infty]$.假定 f 非负,$(A_k, 1\leqslant k\leqslant n)$为 Ω 的有限 $\mathscr{F}$-集分割.定义 f 关于 μ 的积分为

$$\int_\Omega f\mathrm{d}\mu = \sup\Big\{\sum_{i=1}^{n}\inf_{\omega\in A_i}\{f(\omega)\}\mu(A_i)\Big\},$$

这里"sup"表示对 Ω 的所有有限 $\mathscr{F}$-集分割求上确界.假定 f 是一般可测函数.令 $f^+=f\vee 0$, $f^-=f^+-f$.则 $f^\pm$ 为非负可测函数.若 $\int f^+\,\mathrm{d}\mu\wedge\int f^-\,\mathrm{d}\mu<\infty$,则定义 f 关于 μ 的积分为

$$\int f\mathrm{d}\mu=\int f^+\,\mathrm{d}\mu-\int f^-\,\mathrm{d}\mu.$$

定理 2.1.1　(i) 若 f 为非负简单函数,即 $f=\sum_{i=1}^{n}x_i 1_{A_i}$,其中 $x_i\in\mathrm{R}$ 且非负,$(A_i, 1\leqslant i\leqslant n)$为 Ω 的有限 $\mathscr{F}$-集分割,则

$$\int f\mathrm{d}\mu=\sum_{i=1}^{n}x_i\mu(A_i).$$

(ii) 若 f,g 可测,且 $0\leqslant f\leqslant g$,则 $\int f\mathrm{d}\mu\leqslant\int g\mathrm{d}\mu$.

(iii) 若 f,f_n 可测,且 $0\leqslant f_n\uparrow f$,则 $\int f_n\mathrm{d}\mu\uparrow\int f\mathrm{d}\mu$.

(iv) 若 f,g 为非负可测函数,α,β 为非负实数,则

$$\int(\alpha f+\beta g)\mathrm{d}\mu=\alpha\int f\mathrm{d}\mu+\beta\int f\mathrm{d}\mu.$$

证 只证(iii). 其它可直接按定义证明. 结论(iii)$\Longleftrightarrow$

$$\lim_{n\to\infty}\int f_n\mathrm{d}\mu\geqslant\int f\mathrm{d}\mu \text{ 或 }\geqslant s=\sum_{i=1}^{n}v_i\mu(A_i),\qquad(2.1\text{-}1)$$

其中$(A_i,1\leqslant i\leqslant n)$为$\Omega$的有限$\mathscr{F}$-集分割,

$$v_i=\inf\{f(\omega):\omega\in A_i\}.$$

设$\varepsilon>0,A_{in}=\{\omega\in A_i:f_n(\omega)>v_i-\varepsilon\}$,则$A_{in}\uparrow A_i(n\uparrow\infty)$ $(A_{in},1\leqslant i\leqslant m)$为不相联的$\mathscr{F}$-集族. 于是

$$\int f_n\mathrm{d}\mu\geqslant\sum_{i=1}^{m}(v_i-\varepsilon)\mu(A_{in})\xrightarrow[n\to\infty]{}\sum_{i=1}^{m}(v_i-\varepsilon)\mu(A_i)$$

$$=s-\varepsilon\sum_{i=1}^{m}\mu(A_i).\qquad(2.1\text{-}2)$$

若$s,v_i,\mu(A_i),1\leqslant i\leqslant m$,均是有限的,则让$\varepsilon\downarrow 0$,即可由(2.1-2)式得(2.1-1)式. 若$s$有限,则$v_i\mu(A_i),1\leqslant i\leqslant m$,均有限. 去掉其中为0的项,不妨就设前$m_0$项不为0. 这时,$\mu(A_i),1\leqslant i\leqslant m_0$,有限. 将(2.1-2)式中的$m$改成$m_0$,并让$\varepsilon\downarrow 0$,则由(2.1-2)式即可得(2.1-1)式. 若$s=\infty$,则存在i_0,使得$v_{i_0}\mu(A_{i_0})=\infty$. 于是,$v_{i_0}$和$\mu(A_{i_0})$中至少有一个为$\infty$. 设$0<x<v_{i_0},0<y<\mu(A_{i_0}),A_{i_0n}=\{\omega\in A_{i_0}:f_n(\omega)>x\}$. 由$f_n\uparrow f$知:$A_{i_0n}\uparrow A_{i_0}$. 由测度的连续性知:对足够大的$n$,有$\mu(A_{i_0n})>y$. 于是,对足够大的$n$,有

$$\int f_n\mathrm{d}\mu\geqslant x\mu(A_{i_0n})\geqslant x\cdot y.$$

如果$v_{i0}=\infty$,就让$x\uparrow\infty$;如果$\mu(A_{i_0})=\infty$,就让$y\uparrow\infty$. 总之有$\sup_n\left(\int f_n\mathrm{d}\mu\right)=\infty\geqslant\sum_{i=1}^{m}v_i\mu(A_i)$,即(2.1-1)式成立.

注1 上述定理表达积分定义2.1.1的合理性. 也就是说:积分不论如何定义都由定理中的四条性质唯一决定.

定义2.1.2 设f为可测函数. 若$\int|f|^p\mathrm{d}\mu<\infty$,则称$f$关于$\mu$是$p$方可积的. $L^p=L^p(\Omega,\mathscr{F},\mu)$表示关于$\mu$是$p$方可积之

函数全体.

定理 2.1.2 设 $0\leqslant f_n\uparrow f(\mathrm{mod}\,\mu)$,则 $\int f_n\mathrm{d}\mu\uparrow\int f\mathrm{d}\mu$.

证 设 $A=\{\omega:\ f_n(\omega)\uparrow f(\omega)\}$,则 $f_n\cdot 1_A\uparrow f1_A$,且 $\mu(A^c)=0$. 按定理 2.1.1(iii),有

$$\int f_n\mathrm{d}\mu=\int f_n1_A\mathrm{d}\mu\uparrow\int f1_A\mathrm{d}\mu=\int f\mathrm{d}\mu.$$

注 2 条件"$0\leqslant f_n$"可减弱为条件"$f_n\geqslant g\in L^1(\Omega,\ \mathscr{F},\ \mu)$". 但 $g\in L^1$ 不能再减弱. 例如,设 $\Omega=[0,1)$,$\mathscr{F}=\Omega\cap\mathscr{B}$,$P$ 为$(\Omega,\mathscr{F})$上的 Lebesque 测度. 令 $x_0=0$, $x_n\uparrow 1$,$\Delta_n=[x_{n-1},\ x_n)$, $|\Delta_n|=x_n-x_{n-1}$. 定义

$$\xi=\sum_{n=1}^{\infty}\frac{1}{n|\Delta_n|}1_{\Delta_n}(\omega),$$

则 ξ 为$(\Omega,\ \mathscr{F},\ P)$上的 r.v.,且 $\int_\Omega\xi\mathrm{d}P=\sum_n\frac{1}{n}=\infty$. 现在,$\eta_n=n\wedge\xi-\xi$,则 $\eta_n\uparrow 0$,且 $\int_\Omega\eta_n\mathrm{d}P\leqslant n-\int_\Omega\xi\mathrm{d}P=-\infty\ (\forall\ n\geqslant 1)$. 因此,$\int_\Omega\eta_n\mathrm{d}P\to-\infty$. 这表明:lim 与 $\int_\Omega$ 不能交换,原因在于破坏了条件"$g\leqslant\eta_n,\ g\in L^1$".

定理 2.1.3 若 $g\leqslant f_n(\forall\ n\geqslant 1)$,且 $g\in L^1$,则

$$\int_\Omega\underline{\lim}f_n\mathrm{d}\mu\leqslant\underline{\lim}\int_\Omega f_n\mathrm{d}\mu.$$

证 不妨设 $g=0$. 令 $g_n=\inf\{f_k:\ k\geqslant n\}$,则

$$0\leqslant g_n\uparrow\lim_{n\to\infty}g_n=\underline{\lim}f_n.$$

按定理 2.1.2,

$$\int_\Omega\underline{\lim}f_n\mathrm{d}\mu\leqslant\lim_{n\to\infty}\int_\Omega g_n\mathrm{d}\mu\leqslant\underline{\lim}\int_\Omega f_n\mathrm{d}\mu.$$

注 3 这是著名的 Fatou 引理,条件"$f_n\geqslant g\in L^1$"也不能减弱. 注 2 中的例同样也可用来证实这一点.

定理 2.1.4(Lebesque 控制收敛定理) 设 $|f_n|\leqslant g\in L^1$, 且 $f_n\to f(\mathrm{mod}\,\mu)$. 则 $f\in L^1$,且

$$\lim_{n\to\infty}\int_\Omega f_n \mathrm{d}\mu = \int_\Omega f\mathrm{d}\mu.$$

证　"$|f_n|\leqslant g$"等价于"$-g\leqslant f_n\leqslant g$"等价于"$-g\leqslant f_n$, $-g\leqslant -f_n$". 按 Fatou 引理, $\int_\Omega \underline{\lim} f_n \mathrm{d}\mu \leqslant \underline{\lim}\int f_n \mathrm{d}\mu$,

$$\int_\Omega \underline{\lim}(-f_n)\mathrm{d}\mu \leqslant \underline{\lim}\int -f_n\mathrm{d}\mu.$$

$$\Rightarrow \int_\Omega f\mathrm{d}\mu = \int_\Omega \underline{\lim} f_n\mathrm{d}\mu \leqslant \overline{\lim}\int f_n\mathrm{d}\mu \leqslant \int_\Omega \overline{\lim} f_n\mathrm{d}\mu = \int_\Omega f\mathrm{d}\mu.$$

推论 2.1.1　设 $\mu(\Omega)<\infty$, $(f_n, n\geqslant 1)$一致有界, 且 $f_n\to f(\mathrm{mod}\,\mu)$, 则 $\int_\Omega f_n\mathrm{d}\mu \to \int_\Omega f\mathrm{d}\mu$.

证　设 K 为$(f_n, n\geqslant 1)$的界. 由 $\mu(\Omega)<\infty$知: K 是可积的. 于是, 按定理 2.1.4, 立即可得此推论.

定理 2.1.5　设$(\Omega, \mathscr{F})$, $(\Omega', \mathscr{F}')$为可测空间. ξ 为 Ω 上的 Ω'-值可测映射. f 为 Ω' 上的非负实值可测函数, 则

$$\int_\Omega f\circ\xi(\omega)\mu(\mathrm{d}\omega) = \int_{\Omega'} f(\omega')\mu\xi^{-1}(\mathrm{d}\omega').$$

其中 μ 为 $\mathscr{F}$ 上的测度, $\mu\xi^{-1}$为 $\mathscr{F}'$ 上的测度:

$$\mu\xi^{-1}(B') = \mu(\xi^{-1}(B')) \quad (\forall\ B'\in\mathscr{F}').$$

证　如果 $f=1_{B'}(B'\in\mathscr{F}')$. 则 $f\circ\xi=1_{\xi^{-1}(B')}$.

$$\Rightarrow \int_\Omega f\circ\xi\mathrm{d}\mu = \mu(\xi^{-1}(B')) = \mu\xi^{-1}(B') = \int_{\Omega'} f\mathrm{d}\mu\xi^{-1}.$$

这表明: 对这种简单的 f 结论成立. 一般情形, 可先构作简单函数列, 然后, 通过极限手续即可达到定理中的一般结论.

定义 2.1.3　设$(f_n, n\geqslant 1)$为定义在$(\Omega, \mathscr{F}, \mu)$上的可积函数列. 若

$$\lim_{\alpha\to\infty}\sup_n\int_{\{|f_n|\geqslant\alpha\}}|f_n|\mathrm{d}\mu = 0,$$

则称$(f_n, n\geqslant 1)$关于 μ 是**一致可积**的.

定理 2.1.6　设 $\mu(\Omega)<\infty$, $f_n\to f(\mathrm{mod}\,\mu)$, 且$(f_n, n\geqslant 1)$一致可积, 则 $f\in L'$, 且

$$\lim_{n\to\infty}\int_\Omega f_n \mathrm{d}\mu = \int_\Omega f \mathrm{d}\mu.$$

证 $\int_\Omega |f_n| \mathrm{d}\mu = \left(\int_{\{|f_n|\leqslant\alpha\}} + \int_{\{|f_n|>\alpha\}}\right)|f_n| \mathrm{d}\mu$

$$\leqslant \alpha\mu(\Omega) + \sup_n\int_{\{|f_n|>\alpha\}}|f_n|\mathrm{d}\mu.$$

$\Rightarrow \int_\Omega |f_n|\mathrm{d}\mu \leqslant \alpha\mu(\Omega) + 1$ (当 α 足够大时).

$\Rightarrow \left(\int |f_n|\mathrm{d}\mu,\ n\geqslant 1\right)$ 有界. $\Rightarrow f\in L^1$ (Fatou 引理).

令 $f_n^{(\alpha)} = f_n 1_{\{|f_n|\leqslant\alpha\}}$, $f^{(\alpha)} = f 1_{\{|f|\leqslant\alpha\}}$. 则

$$\int |f_n - f|\mathrm{d}\mu \leqslant \int |f_n^{(\alpha)} - f^{(\alpha)}|\mathrm{d}\mu$$

$$+ \int |f^{(\alpha)} - f|\mathrm{d}\mu + \int |f_n - f_n^{(\alpha)}|\mathrm{d}\mu.$$

此式右边第一项当 α 固定, $n\to\infty$ 时以 0 为极限(利用 Lebesque 控制收敛定理);第二项当 $\alpha\uparrow\infty$ 时极限为 0(利用 $f\in L^1$);第三项当 $\alpha\uparrow\infty$ 时极限为 0(利用 (f_n) 的一致可积性). 故定理成立.

注 4 上述定理中的条件 $\mu(\Omega)<\infty$ 不能减弱. 例如,设 $\mu(\Omega)=\infty$, $(A_n,\ n\geqslant 1)$ 为 Ω 的 $\mathscr{F}$-集覆盖,且 $A_n\uparrow$, $\mu(A_n)<\infty$ $(\forall\ n\geqslant 1)$. 令 $f_n = 1_{A_n}$,则 $(f_n,\ n\geqslant 1)$ 一致有界,更是一致可积的,且 $f_n\uparrow 1$ $(\mathrm{mod}\,\mu)$. 定理 2.1.6 对此种情形不成立.

定理 2.1.7 设 $\mu(\Omega)<\infty$, $(f_n,\ n\geqslant 1)\subset L^1$,则 $(f_n,\ n\geqslant 1)$ 一致可积的充要条件是:

(i) $\left(\int |f_n|\mathrm{d}\mu,\ n\geqslant 1\right)$ 有界;

(ii) $\sup_n \int_A |f_n|\mathrm{d}\mu \to 0$ 当 $\mu(A)\to 0$ 时.

证 令 $B_\alpha^n = \{|f_n|\geqslant\alpha\}$,则

$$\int_A |f_n|\mathrm{d}\mu = \int_{A\cap B_\alpha^n}|f_n|\mathrm{d}\mu + \int_{A\setminus B_\alpha^n}|f_n|\mathrm{d}\mu$$

$$\leqslant \int_{B_\alpha^n}|f_n|\mathrm{d}\mu + \alpha\mu(A);$$

$$\int_\Omega |f_n|\mathrm{d}\mu \leqslant \alpha\mu(\Omega) + \sup_n \int_{B_\alpha^n} |f_n|\mathrm{d}\mu.$$

由上述二式立即可推出条件的必要性.

下证充分性,按 Chebyshev 不等式,有

$$\sup_n \mu(B_\alpha^n) \leqslant \frac{1}{\alpha} \sup_n \int_\Omega |f_n|\mathrm{d}\mu \xrightarrow[\alpha\to\infty]{} 0 \ (\text{由条件(i)})$$

于是,按条件(ii),对$\forall\ \varepsilon>0$,$\exists\ \alpha_0=\alpha(\varepsilon)>0$,$\ni$ 当 $\alpha\geqslant\alpha_0$ 时,

$$\sup_n \int_{B_\alpha^m} |f_n|\mathrm{d}\mu < \varepsilon \quad (\forall\ m\geqslant 1)$$

$$\Rightarrow \int_{B_\alpha^n} |f_n|\mathrm{d}\mu \leqslant \sup_k \int_{B_\alpha^n} |f_k|\mathrm{d}\mu < \varepsilon \quad (\forall\ n\geqslant 1). \ \text{当}\ \alpha\geqslant\alpha_0\ \text{时},$$

$$\Rightarrow \sup_n \int_{B_\alpha^n} |f_n|\mathrm{d}\mu \leqslant \varepsilon \quad \text{当}\ \alpha\geqslant\alpha_0\ \text{时},$$

$$\Rightarrow \lim_{\alpha\to\infty} \sup_n \int_{B_\alpha^n} |f_n|\mathrm{d}\mu = 0.$$

定理 2.1.8 设$(f_n,\ n\geqslant 1)\subset L'$,$G(t)$为 $\mathrm{R}_+=[0,\infty)$上的非负增函数,且 $\lim\limits_{t\to\infty}\dfrac{G(t)}{t}=\infty$,$\sup\limits_n\int_\Omega G(|f_n|)\mathrm{d}\mu<\infty$,则 $(f_n,\ n\geqslant 1)$一致可积.

证 当 α 足够大时,有

$$\int_{\{|f_n|\geqslant\alpha\}} |f_n|\mathrm{d}\mu = \int_{\{|f_n|\geqslant\alpha\}} |f_n|\cdot G^{-1}(|f_n|)\cdot G(|f_n|)\mathrm{d}\mu$$

$$\leqslant \varepsilon\cdot\sup_n \int_\Omega G(|f_n|)\mathrm{d}\mu, \quad \text{当}\ \alpha\geqslant t_0\ \text{时},$$

其中 t_0 满足 $\sup\left\{\dfrac{t}{G(t)},\ t\geqslant t_0\right\}\leqslant\varepsilon$ (按条件,t_0 存在).

$$\Rightarrow \sup_n \int_{\{|f_n|\geqslant\alpha\}} |f_n|\mathrm{d}\mu \leqslant \varepsilon \sup_n \int_\Omega G(|f_n|)\mathrm{d}\mu \ (\alpha\geqslant t_0).$$

故$(f_n,\ n\geqslant 1)$一致可积.

注 5 一致可积性概念不涉及 $\mu(\Omega)$是否有限的问题,但真正使人们感兴趣的则是 $\mu(\Omega)<\infty$. 然而,这个限制对结论并没有改善. 下面是这个结论的一个重要的特例.

推论 2.1.2 设$(f_n, n\geqslant 1)\in L^1$,若$\exists\ \nu>0$,使得

$$\left(\int_\Omega |f_n|^{1+\nu}\mathrm{d}\mu,\ n\geqslant 1\right)\text{有界},$$

则$(f_n, n\geqslant 1)$一致可积.

定理 2.1.9(Jensen 不等式) 设 g 为 R→R 的连续下凸函数. $f\in L^1$, $\mu(\Omega)=1$. 则

$$g(Ef)\leqslant Eg(f),\ E\xi=\int_\Omega \xi\mathrm{d}\mu.$$

证 由下凸性知:对$\forall\ x_0\in\mathrm{R}$,存在数 $\lambda(x_0)$,使得

$$g(x)\geqslant g(x_0)+(x-x_0)\lambda(x_0)\quad(\forall\ x\in\mathrm{R}).$$

在此式中,让 $x_0=Ef$, $x=f$,则得

$$g(f)\geqslant g(Ef)+(f-Ef)\lambda(Ef).$$

在此式两边作用算子 E,则因 $E(f-Ef)=0$,而有

$$Eg(f)\geqslant g(Ef).$$

(注意:$\mu(\Omega)=1$ 之条件不可少.)

定理 2.1.10(Lyapunov 不等式) 设对某 $p>0$,有 $f\in L^p$,且 $\mu(\Omega)=1$,则

$$(E|f|^s)^{\frac{1}{s}}\leqslant(E|f|^t)^{\frac{1}{t}}\quad(0<s\leqslant t\leqslant p).$$

证 令 $r=\dfrac{t}{s}>1$(等于 1 的情形是平凡的),$\eta=|f|^s$, $g(x)=|x|^r$(此函数下凸). 则按定理 2.1.9,得

$$(E\eta)^r\leqslant E\eta^r\Rightarrow(E|f|^s)^{\frac{t}{s}}\leqslant E|f|^t.$$

定理 2.1.11(Hölder 不等式) 设 $1<p<\infty$, $1<q<\infty$,且$\dfrac{1}{p}+\dfrac{1}{q}=1$,则当 $f\in L^q$, $g\in L^p$ 时,有

$$E_\mu|f\cdot g|\leqslant(E_\mu|f|^q)^{\frac{1}{q}}(E_\mu|g|^p)^{\frac{1}{p}}\quad\left(E_\mu\xi=\int_\Omega\xi\mathrm{d}\mu\right).$$

证 按对数函数的上凸性,有

$$\ln(ax+by)\geqslant a\ln x+b\ln y\ (x,y,a,b\in R_+,\ a+b=1)$$

即

$$\ln(ax+by)\geqslant\ln x^a y^b.$$

令 $x=|g|^p(E_\mu|g|^p)^{-1}$，$y=|f|^q(E_\mu|f|^q)^{-1}$，$a=\dfrac{1}{p}$，$b=\dfrac{1}{q}$，则 $E_\mu(x^{\frac{1}{p}}\cdot y^{\frac{1}{q}})\leqslant 1$. 代真后即得所要的不等式.

定理 2.1.12(Minkowski 不等式) 若 $f,g\in L^p(1\leqslant p<\infty)$，则 $E_\mu|f+g|^p<\infty$，且

$$(E_\mu|f+g|^p)^{\frac{1}{p}}\leqslant(E_\mu|f|^p)^{\frac{1}{p}}+(E_\mu|g|^p)^{\frac{1}{p}}.$$

证 显然，有

$$|f+g|^p\leqslant 2^p(|f|\vee|g|)^p\leqslant 2^p(|f|^p+|g|^p)$$
$$\Rightarrow f+g\in L^p.$$

当 $p=1$ 时，定理中的不等式显然成立. 下面考虑 $p>1$ 的情形. 这时，有

$$\begin{aligned}|f+g|^p&\leqslant|f+g||f+g|^{p-1}\\&\leqslant|f||f+g|^{p-1}+|g||f+g|^{p-1}.\end{aligned}$$

让 q 满足 $\dfrac{1}{p}+\dfrac{1}{q}=1$，则按 Hölder 不等式，有

$$\begin{aligned}E_\mu|f+g|^p\leqslant&(E_\mu|f|^p)^{\frac{1}{p}}(E_\mu|f+g|^{(p-1)q})^{\frac{1}{q}}\\&+(E_\mu|g|^p)^{\frac{1}{p}}(E_\mu|f+g|^{(p-1)q})^{\frac{1}{q}}\end{aligned}$$

$$\Rightarrow E_\mu|f+g|^p\leqslant(E_\mu|f+g|^p)^{\frac{p-1}{p}}((E_\mu|f|^p)^{\frac{1}{p}}+(E_\mu|f|^p)^{\frac{1}{p}}).$$

由此即得所要的不等式.

定理 2.1.13(Cantelli 不等式) 设$(\Omega,\mathscr{F},P)$为概率空间，X 为定义于其上的可积 r.v.，$m=Ex$，$\sigma^2=E(X-Ex)^2$，$\left(EX=\int_\Omega X\mathrm{d}P\right)$，则

$$P(X-m\geqslant\alpha)\leqslant\frac{\sigma^2}{\sigma^2+\alpha^2}\ (\forall\,\alpha>0).$$

证 不妨设 $m=0$，定义

$$\mu_0(A)=P(A\cap\{x<\alpha\}),\ \mu_1(A)=P(A\cap\{x\geqslant\alpha\})$$

其中 $A\in\mathscr{F}$，则 μ_0 和 μ_1 为$(\Omega,\mathscr{F})$上的测度. 按 Hölder 不等式，有

$$(E_{\mu_0}X)^2 \leqslant \mu_0(\Omega)\cdot E_{\mu_0}X^2,\ (E_{\mu_1}X)^2 \leqslant \mu_1(\Omega)\,E_{\mu_1}X^2.$$

注意下列各式：

$$\begin{aligned}
0 &= EX = E_{\mu_0}X + E_{\mu_1}X;\\
\sigma^2 &= EX^2 = E_{\mu_0}X^2 + E_{\mu_1}X^2\\
&\geqslant \frac{1}{\mu_0(\Omega)}(E_{\mu_0}X)^2 + \frac{1}{\mu_1(\Omega)}(E_{\mu_1}X)^2\\
&= \frac{1}{\mu_0(\Omega)\mu_1(\Omega)}(E_{\mu_1}X)^2 \geqslant a^2\,\frac{\mu_1(\Omega)}{\mu_0(\Omega)}
\end{aligned}$$

$$\Rightarrow \sigma^2\mu_0(\Omega) \geqslant a^2\mu_1(\Omega),\quad \mu_1(\Omega) = P(X \geqslant a),$$

$$\mu_0(\Omega) = P(X < a) = 1 - \mu_1(\Omega)$$

$$\Rightarrow \mu_1(\Omega) \leqslant \frac{\sigma^2}{\sigma^2 + a^2}.$$

定理 2.1.14（分部积分公式）　设 F,G 为 $[a,b]$ 上的非降右连左极函数. μ,ν 为相应的测度：

$$\mu(x,y] = F(y) - F(x),\ \nu(x,y] = G(y) - G(x)$$
$$(a \leqslant x \leqslant y \leqslant b).$$

若 F,G 在 $(a,b]$ 中无共同间断点，则

$$\int_{(a,b]} G(x)\mu(\mathrm{d}x) = F(b)G(b) - F(a)G(a) - \int_{(a,b]} F(x)\nu(\mathrm{d}x).$$

证　假设条件用于保证：

$$\int_{(a,b]} \nu(\{x\})\mu(\mathrm{d}x) = 0 = \int_{(a,b]} \mu(\{x\})\nu(\mathrm{d}x).$$

§2.2　随机变量列的收敛类型

本节如无特别说明，所考虑的概率空间均指$(\Omega, \mathscr{F}, P)$.

定义 2.2.1　若 $A,B\in\mathscr{F}$，且 $P(A\cap B)=P(A)P(B)$，则说 A 和 B 关于 P 独立（以下将省去“关于 P”一词），若 $(A_k,\ 1\leqslant k\leqslant n)\subset\mathscr{F}$，且

$$P(\bigcap_{i=1}^{j} A_{k_i}) = \prod_{i=1}^{j} P(A_{k_i})$$

$$(1 \leqslant j \leqslant n,\ (k_1, \cdots, k_j) \subset (1, 2, \cdots, n)),$$

则说$(A_k,\ 1 \leqslant k \leqslant n)$是独立的. 若$(A_n,\ n \geqslant 1) \subset \mathscr{F}$,且

$$P(\bigcap_{i=1}^{n} A_{k_i}) = \prod_{i=1}^{n} P(A_{k_i})$$

$$(n \geqslant 1,\ (k_i,\ 1 \leqslant i \leqslant n) \subset \mathrm{N} = \{1, 2, \cdots\}),$$

则说$(A_n,\ n \geqslant 1)$是独立的. 若$\mathscr{A}_n \subset \mathscr{F}$,且$\Omega \in \mathscr{A}_n\ (n \geqslant 1)$,并

$$P(\bigcap_{k=1}^{n} A_k) = \prod_{k=1}^{n} P(A_k)\ (\forall\ n \in N,\ A_k \in \mathscr{A}_k,\ 1 \leqslant k \leqslant n),$$

则说$(\mathscr{A}_n,\ n \geqslant 1)$独立. 若$(x_n,\ n \geqslant 1)$为r.v.列,且

$(\sigma(x_n),\ n \geqslant 1)$独立,则说$(x_n,\ n \geqslant 1)$为独立r.v.列.

按此定义,元素$\varnothing$,Ω在任何概率测度P下都与$\mathscr{F}$独立. 如果$A \in \mathscr{F}$与自身独立,则$A = \varnothing$或Ω.

定理 2.2.1(Borel-Cantelli 引理) 设$(A_n,\ n \geqslant 1) \subset \mathscr{F}$.

(i) 若$\sum_n P(A_n) < \infty$,则$P(\overline{\lim} A_n) = 0$.

(ii) 若$\sum_n P(A_n) = \infty$且$(A_n,\ n \geqslant 1)$独立,则$P(\overline{\lim} A_n) = 1$.

证 $\overline{\lim} A_n = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k$.

若$\sum_n P(A_n) < \infty$,则$P(\bigcup_{k=n}^{\infty} A_k) \leqslant \sum_{k=n}^{\infty} P(A_k) \xrightarrow[n \to \infty]{} 0 \Rightarrow$(i).

若$\sum_n P(A_n) = \infty$, 则

$$\begin{aligned} P(\bigcap_{k=n}^{n+m} A_k^c) &= \prod_{k=n}^{n+m} P(A_k^c) = \prod_{k=n}^{n+m} (1 - P(A_k)) \\ &\leqslant \prod_{k=n}^{n+m} \exp(-P(A_k)) \\ &= \exp\Big(-\sum_{k=n}^{n+m} P(A_k)\Big) \xrightarrow[m \to \infty]{} 0\ (\forall\ n \geqslant 1) \end{aligned}$$

$$\Rightarrow P(\bigcap_{k=n}^{\infty} A_k^c) = 0 \ (\forall\ n \geqslant 1) \Rightarrow \text{(ii)}.$$

定理 2.2.2(Kolmogorov 0-1 律) 设$(A_n,\ n\geqslant 1)\subset\mathscr{F}$，且独立. $\mathscr{I}=\bigcap_{n=1}^{\infty}\sigma(A_n,\ A_{n+1},\cdots)$ (一般称为尾 σ-代数).
则 $P(A)=0$ 或 $1(\forall\ A\in\mathscr{I})$.

证 $A\in\mathscr{I}\Rightarrow A\in\sigma(A_n,\ A_{n+1},\cdots)\ (\forall\ n\geqslant 1)$. 于是，按后面的引理 2.2.1，$A$ 与$(A_k,\ 1\leqslant k\leqslant n-1)$独立$(\forall\ n\geqslant 1)$. 由归纳法立即可得 A 与$(A_n,\ n\geqslant 1)$独立. 按引理 2.2.1，A 与 $\sigma(A_n,\ n\geqslant 1)$独立. 然而，$A\in\sigma(A_n,\ n\geqslant 1)$. 故 A 与 A 独立. 从而 $P(A)=P(A\cdot A)=P^2(A)$.

$\Rightarrow P(A)=0$ 或 1.

引理 2.2.1 设 $\mathscr{A}=(A_n,\ n\geqslant 1)$. 若

(i) $A\in\mathscr{F}$ 独立于 $\mathscr{A}$，且$(A_n,\ n\geqslant 1)$不相联，

或 (ii) $\{A,\ \mathscr{A}\}$独立；

则 A 独立于 $\sigma(\mathscr{A})$.

证 首先设条件(i)成立. 令

$$\Omega_0 = \bigcup_{n=1}^{\infty} A_n,\ A_0 = \Omega_0^c,$$

则 A 与 Ω_0 独立. 从而，与 A_0 独立. 令 $\mathscr{A}^*=(A_n,\ n\geqslant 0)$，则 $\mathscr{A}^*$ 为 Ω 的一个分割，且 A 与 $\mathscr{A}^*$ 独立. 显然，$\sigma(\mathscr{A})=\sigma(\mathscr{A}^*)$. 于是，结论等价于“$A$ 独立于 $\sigma(\mathscr{A}^*)$”. 易证，$\sigma(\mathscr{A}^*)=\{A^*：A^*=\bigcup_{i=1}^{\infty}A_{k_i},\ (k_1,k_2,\cdots)\subset\mathrm{N},\ A_{k_i}\in\mathscr{A}^*\}$.

注意：$(A_{k_i},\ i\geqslant 1)$不相联. 因此，$A^*=\bigcup_{i=1}^{\infty}A_{k_i}\in\sigma(\mathscr{A}^*)$与 A 独立. 故 A 独立于 $\sigma(\mathscr{A}^*)$.

现在假定条件(ii)成立. 令

$$B_1 = A_1,\ B_n = A_n\backslash\bigcup_{k=1}^{n-1}A_k \quad (n \geqslant 2),$$

则 $\tilde{\mathscr{A}}=(B_n,\ n\geqslant 0)$为 Ω 的一个分割，其中 $B_0=(\bigcup_{n=1}^{\infty}B_n)^c$，且 $\sigma(\mathscr{A})=\sigma(\tilde{\mathscr{A}})$，$A$ 与 $\tilde{\mathscr{A}}$独立. 故 A 独立于 $\sigma(\mathscr{A})$.

定义 2.2.1 设(ξ_∞, ξ_n, $n\geqslant1$)是定义在(Ω, $\mathscr{F}$, P)上的 r.v. 族. 若对$\forall$ $\varepsilon>0$, 有$\lim\limits_{n\to\infty}P(|\xi_n-\xi_\infty|>\varepsilon)=0$, 则称$\xi$**依概率** P **收敛于**ξ_∞. 记为$\xi_n\xrightarrow{P.}\xi_\infty$或$P-\lim\limits_{n\to\infty}\xi_n=\xi_\infty$. 若$\lim\limits_{n,m\to\infty}P(|\xi_n-\xi_m|>\varepsilon)=0$, 则称$\xi$为$P$-基本列. 若$P(\omega:\ \xi_n\not\to\xi_\infty)=0$, 则说$\xi$(a.s) **收敛于**$\xi_\infty$, 记为$\xi_n\xrightarrow{\text{a.s}}\xi_\infty$或 (a.s)$-\lim\limits_{n\to\infty}\xi_n=\xi_\infty$. 若$P(\omega:\ \xi$不收敛$)=0$, 则说$\xi$为$P$-a.s 基本列. 若对某$p>0$, 有

$$E|\xi_n-\xi_\infty|^p\to0\ (\lim_{n,m\to\infty}E|\xi_n-\xi_m|^p=0),$$

则说ξ是 p **方收敛于**ξ_∞(p 方基本列). 记为$\xi_n\xrightarrow{L^p}\xi_\infty$或$L^p-\lim\limits_{n\to\infty}\xi_n=\xi_\infty$. 若对任意有界连续函数$f$, 有$Ef(\xi_n)\to Ef(\xi_\infty)$, 则说$\xi$ **依分布收敛于**ξ_∞. 记为$\xi_n\xrightarrow{d.}\xi_\infty$或$d-\lim\limits_{n\to\infty}\xi_n=\xi_\infty$.

注 1 依概率收敛只涉及到变量族的二维分布. 而P-a.s 收敛(或说以概率 1 收敛)却涉及到变量族的整个有限维分布族. 由此可见, P-a.s 收敛必导致P-收敛.

注 2 易证: "$\xi_n\xrightarrow{d.}\xi_\infty$"等价于"对$\forall$ $\lambda\in\mathbf{R}$, 有

$$E\exp(i\lambda\xi_n)\to E\exp(i\lambda\xi_\infty)".$$

定理 2.2.3 ξ以概率 1 收敛等价于下述条件:

$$\lim_{n\to\infty}P(\sup\{|\xi_k-\xi_\infty|,\ k\geqslant n\}\geqslant\varepsilon)=0\ (\forall\ \varepsilon>0).$$

证 令$A_n^\varepsilon=\{\omega:\ |\xi_n-\xi_\infty|\geqslant\varepsilon\}$.

$$A^\varepsilon=\overline{\lim_{n\to\infty}}A_n^\varepsilon=\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty}A_k^\varepsilon.$$

则当$\varepsilon\downarrow0$时$A^\varepsilon\uparrow$, 且

$$\{\omega:\ \xi_n\not\to\xi_\infty\}=\bigcup_{\varepsilon>0}A^\varepsilon=\bigcup_{m=1}^{\infty}A^{\frac{1}{m}}.$$

事实上, $\xi_n(\omega)\not\to\xi_\infty(\omega)$表明: 对$\forall$ $\varepsilon>0$及$n\geqslant1$, $\exists$ $k_0\geqslant n$, 使 $|\xi_{k_0}(\omega)-\xi_\infty(\omega)|\geqslant\varepsilon$. 因此, 对$\forall$ $n\geqslant1$, 有$\omega\in A_{k_0}^\varepsilon\subset\bigcup\limits_{k=n}^{\infty}A_k^\varepsilon$.

$$\Rightarrow\omega\in A^\varepsilon, \text{即}\ \{\omega:\ \xi_n(\omega)\not\to\xi_\infty(\omega)\}\subset\bigcup_{\varepsilon>0}A^\varepsilon.$$

另一方面，$\omega\in\bigcup_{\varepsilon>0}A^{\varepsilon}$ 表明：$\exists\ \varepsilon_0>0$，$\ni\ \omega\in A^{\varepsilon_0}$，即 $\omega\in\bigcup_{k=n}^{\infty}A_k^{\varepsilon_0}(\forall\ n\geqslant 1)$. 于是，对$\forall\ n\geqslant 1$，$\exists\ k_0\geqslant n$，$\ni\ \omega\in A_{k_0}^{\varepsilon_0}$.

$\Rightarrow \xi_n(\omega)\not\rightarrow\xi_\infty(\omega)\Rightarrow\bigcup_{\varepsilon>0}A^{\varepsilon}\subset\{\omega:\ \xi_n\not\rightarrow\xi_\infty\}$.

综合起来，得

$$\bigcup_{\varepsilon>0}A^{\varepsilon}=\{\omega:\ \xi_n(\omega)\not\rightarrow\xi_\infty(\omega)\}.$$

按测度的连续性，

$$P(A^{\varepsilon})=\lim_{n\to\infty}P\Big(\bigcup_{k=n}^{\infty}A_k^{\varepsilon}\Big)$$

$$\Rightarrow P(\xi_n\not\rightarrow\xi_\infty)=P\Big(\bigcup_{m=1}^{\infty}A^{\frac{1}{m}}\Big)=\lim_{m\to\infty}P(A^{\frac{1}{m}})$$

$$=\lim_{m\to\infty}\lim_{n\to\infty}P\Big(\bigcup_{k=n}^{\infty}A_k^{\frac{1}{m}}\Big)$$

$$=\lim_{m\to\infty}\lim_{n\to\infty}P\Big(\sup\{|\xi_k-\xi_\infty|:\ k\geqslant n\}\geqslant\frac{1}{m}\Big).\quad(2.2\text{-}1)$$

由此等式即得所要的结论.

推论 2.2.1 若$\sum_n P(|\xi_n-\xi_\infty|\geqslant\varepsilon)<\infty\ (\forall\ \varepsilon>0)$，则

$$\xi_n\xrightarrow[n\to\infty]{}\xi_\infty(\text{a.s}).$$

证 事实上，

$$P\Big(\sup\{|\xi_k-\xi_\infty|:\ k\geqslant n\}\geqslant\frac{1}{m}\Big)$$

$$\leqslant\sum_{k=n}^{\infty}P\Big(|\xi_k-\xi_\infty|\geqslant\frac{1}{m}\Big).$$

按假设及(2.2-1)式，即得这里的结论.

推论 2.2.2 若$\exists$ 一个正数列$(\varepsilon_n,\ n\geqslant 1)$，且 $\varepsilon_n\downarrow 0$，$\ni$ $\sum_n P(|\xi_n-\xi_\infty|\geqslant\varepsilon_n)<\infty$，则 $\xi_n\xrightarrow[n\to\infty]{}\xi_\infty(\text{a.s})$.

证 令 $A_n=\{|\xi_n-\xi_\infty|\geqslant\varepsilon_n\}$，则按定理 2.2.1(i)，$P(\overline{\lim}A_n)=0$. 注意，对$\forall\ \varepsilon>0$，$\exists\ n\geqslant 1$，使得

$$\bigcup_{k=n}^{\infty}\{|\xi_k-\xi_\infty|\geqslant\varepsilon\}\subset\bigcup_{k=n}^{\infty}\{|\xi_k-\xi_\infty|\geqslant\varepsilon_k\}.$$

$$\Rightarrow \{\omega: \xi_n(\omega) \not\to \xi_\infty(\omega)\} \subset \bigcap_{k=1}^{\infty} \bigcup_{k=n}^{\infty} \{|\xi_k - \xi_\infty| \geqslant \varepsilon_k\} = \overline{\lim} A_n$$

$$\Rightarrow P(\xi_n \not\to \xi_\infty) = 0.$$

定理 2.2.3 下列蕴含关系成立.

(i) $\xi_n \xrightarrow{\text{a.s}} \xi_\infty \Rightarrow \xi_n \xrightarrow{p.} \xi_\infty \Rightarrow \xi_n \xrightarrow{d.} \xi_\infty$;

(ii) $\xi_n \xrightarrow{L^p} \xi_\infty (p>0) \Rightarrow \xi_n \xrightarrow{P.} \xi_\infty$.

证 首先设 $\xi_n \xrightarrow{\text{a.s}} \xi_\infty$，按定理 2.2.2 之证，有

$$\{\omega: \xi_n \not\to \xi_\infty\} = \bigcup_{\varepsilon>0} \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} \{|\xi_k - \xi_\infty| \geqslant \varepsilon\}$$

$$\Rightarrow \{\xi_n \not\to \xi_\infty\} \supseteq \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} \{|\xi_k - \xi_\infty| \geqslant \varepsilon\} \quad (\forall\ \varepsilon > 0)$$

$$\Rightarrow 0 = P(\xi_n \not\to \xi_\infty) \geqslant \lim_{n\to\infty} P\Big(\bigcup_{k=n}^{\infty} \{|\xi_k - \xi_\infty| \geqslant \varepsilon\}\Big)$$

$$\geqslant \lim_{n\to\infty} \sup_{k\geqslant n} P(|\xi_k - \xi_\infty| \geqslant \varepsilon). \quad (\forall\ \varepsilon > 0)$$

$$\Rightarrow P(|\xi_n - \xi_\infty| \geqslant \varepsilon) \xrightarrow[n\to\infty]{} 0 \ (\forall\ \varepsilon > 0)，\text{即}\ \xi_n \xrightarrow{P.} \xi_\infty.$$

现设 $\xi_n \xrightarrow{P.} \xi_\infty$，$f$ 为 R 上的任意有界连续函数，其界为 C，则对$\forall\ \varepsilon>0$ 及 $N>0$，$\exists\ \delta=\delta(\varepsilon,N)>0$，使得

$$|f(x) - f(y)| < \varepsilon \ (|x| \leqslant N,\ |x-y| \leqslant \delta,\ x,y \in \mathrm{R}).$$

注意，ξ_∞为 r. v. 表明：$\exists\ N_0=N_0(\varepsilon)>0$，使得

$$P(|\xi_\infty| > N_0) < \varepsilon.$$

取 $\delta_0=\delta(\varepsilon, N_0(\varepsilon))>0$，则对$\forall\ n\geqslant 1$，有

$$\begin{aligned} &E|f(\xi_n) - f(\xi_\infty)| \\ &= E|f(\xi_n) - f(\xi_\infty)|[1_{\{|\xi_n-\xi_\infty|>\delta_0\}} \\ &\quad + 1_{\{|\xi_\infty|\leqslant N_0,\ |\xi_n-\xi_\infty|\leqslant\delta_0\}} + 1_{\{|\xi_\infty|>N_0,\ |\xi_n-\xi_\infty|\leqslant\delta_0\}}] \\ &\leqslant \varepsilon + 2C\varepsilon + 2CP(|\xi_n - \xi_\infty| > \delta_0). \end{aligned}$$

$$\Rightarrow E|f(\xi_n) - f(\xi_\infty)| \xrightarrow[n\to\infty]{} 0,\ \text{即}\ \xi_n \xrightarrow{d.} \xi_\infty.$$

最后设 $\xi_n \xrightarrow{L^p.} \xi_\infty$. 则对$\forall\ \varepsilon>0$，$\exists\ n_0=n_0(\varepsilon)\geqslant 1$，使得

$$E|\xi_n - \xi_\infty|^p < \varepsilon \ (\forall\, n \geqslant n_0).$$

因此,对$\forall\ \delta>0$,当$n\geqslant n_0$时,有

$$P(|\xi_n - \xi_\infty| \geqslant \delta) \leqslant \delta^{-p} E|\xi_n - \xi_\infty|^p \leqslant \varepsilon \cdot \delta^{-p},$$ 即

$$\xi_n \xrightarrow{P.} \xi_\infty.$$

注 2 上述定理之逆一般不成立,下面举例说明. 设 $\Omega=[0,1]$,$\mathscr{F}=\mathscr{B}\cap[0,1]$,$P$ 为 Lebesque 测度.

(1) 设 $A_{ni}=\left[\frac{i-1}{n},\ \frac{i}{n}\right]$,$\xi_n^i=1_{A_{ni}}(1\leqslant i\leqslant n<\infty)$. 定义

$$\xi_m = \xi_n^i,\ m = \sum_{k=0}^{n-1} k + i,\ n \geqslant 1,\ 1 \leqslant i \leqslant n.$$

则 $\xi=(\xi_m,\ m\geqslant 1)=\{\xi_1^1;\ \xi_2^1,\ \xi_2^2;\ \xi_3^1,\ \xi_3^2,\ \xi_3^3;\ \cdots\}$.

易证:$\xi_m \xrightarrow{P.} 0$. 但 ξ 不是(a. s)收敛的.

(2) 定义 $\xi_n=\begin{cases} e^n, & \omega\in\left[0,\dfrac{1}{n}\right], \\ 0, & \omega\in\left[\dfrac{1}{n},1\right]. \end{cases}$ 则易证:$\xi_n \xrightarrow{a.s} 0$. 但对$\forall\ p>0$,$(\xi_n,\ n\geqslant 1)$却不是 L^p-收敛的.

定理 2.2.5 设 $\xi=(\xi_n,\ n\geqslant 1)$为非负 r. v. 列. 若下列条件之一成立:

(i) $\xi_n \xrightarrow{a.s} \xi_\infty$,$E\xi_\infty<\infty$,且 $E\xi_n \longrightarrow E\xi_\infty$;

(ii) $\xi_n \xrightarrow{P.} \xi_\infty$,$\xi$ 一致可积,

则 $\xi_n \xrightarrow{L^r.} \xi_\infty$.

证 设条件(i)成立.

令 $\eta_n=\xi_\infty-\xi_n$,则 $\eta_n \xrightarrow{a.s} 0$. 且 $E\eta_n \longrightarrow 0$. 显然,

$$\begin{aligned} E|\xi_n - \xi_\infty| &= E|\xi_n - \xi_\infty|(1_{\{\xi_\infty-\xi_n\geqslant 0\}} + 1_{\{\xi_\infty-\xi_n<0\}}) \\ &= E(\xi_n - \xi_\infty) + 2E(\eta_n \cdot 1_{\{\eta_n>0\}}). \end{aligned}$$

由 $0\leqslant \eta_n 1_{\{\eta_n>0\}} \leqslant \xi_\infty \in L^1$ 知

$$E(\eta_n 1_{\{\eta_n>0\}}) \longrightarrow 0 \text{(按 Lebesque 控制收敛定理)}.$$

$\Rightarrow E|\xi_n-\xi_\infty|\longrightarrow 0$ (按假设 $E(\xi_n-\xi_\infty)\longrightarrow 0$).

现设条件(ii)成立. 设 $\varepsilon>0$, $N>0$,则

$$E|\eta_n| = E[|\eta_n|(1_{\{|\eta_n|\leqslant\varepsilon\}} + 1_{\{\varepsilon<|\eta_n|\leqslant N\}} + 1_{\{N<|\eta_n|\}})]$$

$$\leqslant \varepsilon + NP(\varepsilon < |\eta_n|) + E(|\eta_n|1_{\{|\eta_n|>N\}}).$$

由 $\xi_n \xrightarrow{P.} \xi_\infty$ 知: $P(\varepsilon<|\eta_n|)\longrightarrow 0$.

按 ξ 一致可积,有 $E(|\eta_n|1_{\{|\eta_n|>N\}})\longrightarrow 0$.

$\Rightarrow E|\eta_n|\longrightarrow 0$.

(注:第(ii)条成立可以不必要求 ξ 非负).

定理 2.2.6 $\xi=(\xi_n, n\geqslant 1)$ 为 P-基本列的充要条件是 ξ 中的任一无穷子列必可选出一个(a.s)-基本列.

证 先证充分性. 反证法. 假定 ξ 不是 P-基本列. 则按定义 2.2.2,$\exists\ \varepsilon>0$,使得

$$\overline{\lim_{n,m\to\infty}} P(|\xi_n - \xi_m| \geqslant \varepsilon) = \delta > 0.$$

不妨假定 $P(|\xi_n-\xi_m|\geqslant\varepsilon)\xrightarrow[n,m\to\infty]{}\delta$(如有必要可以通过选子列保证). 按条件,可选一子列 $(\xi_{n_k}, k\geqslant 1)$,使得 $(\xi_{n_k}, k\geqslant 1)$ 为(a.s)-基本列. 但按定理 2.2.4,它必为 P-基本列. 这与 $\delta>0$ 矛盾. 故 ξ 必为 P-基本列.

现证必要性. 考虑到 P-基本列的子列仍为 P-基本列. 不妨设子列就是 ξ 本身,令 $n_1=1$,定义 $n_{k+1}=\inf\{n>n_k: P(|\xi_n-\xi_{n_k}|>2^{-k})<2^{-k}\}$ $(k\geqslant 1)$.

则 $$\sum_{k=1}^{\infty}P(|\xi_{n_{k+1}} - \xi_{n_k}| > 2^{-k}) < \sum_{k=1}^{\infty}2^{-k} = 1 < \infty$$

$\Rightarrow P(\overline{\lim}\{|\xi_{n_{k+1}}-\xi_{n_k}|>2^{-k}\})=0$ (按定理 2.2.1(i)).

令 $$A_k=\{|\xi_{n_{k+1}}-\xi_{n_k}|>2^{-k}\},$$

$$\mathscr{N}=\overline{\lim}A_n=\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty}A_k,$$

则 $\{(\xi_{n_k}, k\geqslant 1)$不收敛$\}\subset$N.

$\Rightarrow(\xi_{n_k}, k\geqslant 1)$为(a.s)-基本列.

§2.3 分布族的弱收敛性

在上节所讨论的四种收敛性概念中，按定理 2.2.4，d-收敛是最弱的. 本节介绍这类弱收敛性的有关结果.

定义 2.3.1 设 F, F_n, $n\geqslant 1$, 均为 R 上的分布函数. 若 $F_n(x)\longrightarrow F(x)$，其中

$$x\in C(F)=\{x\in \mathrm{R}: x \text{ 为 } F \text{ 的连续点}\}.$$

则说 $(F_n,\ n\geqslant 1)$ 概(或弱)收敛到 F. 记为 $F_n\Rightarrow F$ 或 $F_n\xrightarrow{\text{i.g}}F$.

注 1 分布函数 F(定义在 R 上)是有界的、单调不降的，且右连左极的函数. 易证它的不连续点集 $\mathrm{R}\backslash C(F)$ 至多可数.

注 2 在弱收敛概念中要求收敛出现在 F 的连续点处，其原因可由下例解释. 设 $x=(x_n,\ n\geqslant 1)$ 为独立同分布的 r. v. 列，且 $P(x_1=1)=\frac{1}{2}=P(x_1=-1)$. $S_n=\sum_{k=1}^{n}x_k$. 按弱大数律，有

$$P\left(\left|\frac{1}{n}S_n\right|>\varepsilon\right)\to 0\ (\forall\ \varepsilon>0).$$

设 F_n 为 $\frac{1}{n}S_n$ 的分布函数. 当 $x>0$ 时，有

$$F_n(x)=1-P\left(\frac{1}{n}S_n>x\right)\xrightarrow[n\to\infty]{}1.$$

当 $x<0$ 时，有

$$F_n(x)\leqslant P\left(\frac{1}{n}|S_n|\geqslant|x|\right)\xrightarrow[n\to\infty]{}0.$$

$$\Rightarrow F_n(x)\xrightarrow[n\to\infty]{}\Delta(x)=\begin{cases}1, & x\geqslant 0,\\ 0, & x<0.\end{cases}\quad(\forall\ x\neq 0).$$

显然，$\Delta(x)$ 为分布函数，且 $x=0$ 为它的不连续点. 因此，$F_n\Rightarrow\Delta$. 然而，$(F_n(0),\ n\geqslant 1)$ 不收敛于 $\Delta(0)=1$. 事实上，设 n 取奇数，则 $\{S_n=0\}$ 为不可能事件. 从而，$P(S_n\leqslant 0)=P(S_n\geqslant 0)=\frac{1}{2}$，

$$F_n(0) = P\left(\frac{1}{n}S_n \leqslant 0\right) = P(S_n \leqslant 0) = P(S_n < 0) = \frac{1}{2}.$$

$\Rightarrow F_n(0) \to \frac{1}{2}$，其中 n 按奇数趋于 ∞.

这表明，即使$(F_n(0),\ n\geqslant 1)$收敛也不以 1 为极限.

定义 2.3.2 设$(E,\mathscr{E},\rho)$为一尺度空间，其中 ρ 为 E 上的尺度，$\mathscr{E}$ 是由关于尺度 ρ 的开集类所生成的 σ-代数. P，P_n，$n\geqslant 1$，均为$(E,\mathscr{E})$上的概率测度. 如果对$\forall\ A\in\mathscr{E}$，$P(\partial A)=0$，有

$$P_n(A) \to P(A) \quad (\partial A = [A] \cap [A^c],\ [A] \text{ 表示 } A \text{ 的闭包}).$$

则称 P_n 概收敛到 P. 记为 $P_n \Rightarrow P$ 或 $P_n \xrightarrow{i.g} P$.

如果对 E 上的任意有界连续函数 f，有

$$\int_E f(x) P_n(\mathrm{d}x) \to \int_E f(x) P(\mathrm{d}x).$$

则称 P_n 弱收敛到 P. 记为 $P_n \xrightarrow{w.} P$.

注 3 此定义是定义 2.3.1 的概率化. 从定义中可以看出，空间 E 上具有拓扑结构是这里的基本要求. 下面的定理将要讨论概、弱收敛的等价性. 这一点可从下例中得到启发. 设 $\xi=(\xi_n,\ n\geqslant 1)$为独立同分布的 r. v. 列，且 $P(\xi_1=1)=p$，$P(\xi_1=0)=q=1-p$，$S_n=\sum_{k=1}^{n}\xi_k$. 则按弱大数律，$\frac{1}{n}S_n \xrightarrow{P.} p$. 令 $F_n(x)=P\left(\frac{1}{n}S_n\leqslant x\right)$，则

$$F_n(x) \to F(x) = \begin{cases} 1, & x \geqslant p, \\ 0, & x < p. \end{cases} \quad (\text{当 } x \neq p \text{ 时}).$$

因此，$F_n\Rightarrow F$. 另一方面，按 Lebesque 控制收敛定理，

$$E\exp\left(i\lambda \frac{1}{n}S_n\right) \to \exp(i\lambda p) \ (\forall\ \lambda \in \mathrm{R}).$$

因此，$P_n \xrightarrow{w.} P$. 故在此例中，概、弱收敛等价.

定理 2.3.1 下列命题是等价的：

(i) $P_n \xrightarrow{w.} P$；

(ii) $\overline{\lim\limits_{n\to\infty}} P_n(A) \leqslant P(A)$ ($\forall\ A \in \mathscr{E}$，且 A 为闭集)；

(iii) $\underline{\lim\limits_{n\to\infty}} P_n(A) \geqslant P(A)$ ($\forall\ A \in \mathscr{E}$，且 A 为开集)；

(iv) $P_n \Rightarrow P$.

证 显然，(ii)等价于(iii)(因为开集的补为闭集). 因此，只有下列三种情形需要讨论.

先证(i)⇒(ii). 事实上，对$\forall\ \varepsilon>0$，定义

$$f_\varepsilon(x) = g\left(\frac{1}{\varepsilon}\rho(x,A)\right),$$

$$g(t) = 1_{(-\infty,0]}(t) + (1-t)1_{(0,1]}(t) + 0 \cdot 1_{(1,\infty)}(t) \quad (t \in R),$$

$$\rho(x,A) = \inf\{\rho(x,y):\ y \in A\}.$$

令 $A_\varepsilon = \{x:\ \rho(x,A) < \varepsilon\}$，则 $A_\varepsilon \downarrow A$ 当 $\varepsilon \downarrow 0$ 时，

$$f_\varepsilon(x) \downarrow 1_A(x) \text{ 当 } \varepsilon \downarrow 0, \text{且 } x \notin \partial A \text{ 时}.$$

假定 A 为闭集，且(i)成立，则

$$\overline{\lim_{n\to\infty}}\int_E f_\varepsilon(x)P_n(\mathrm{d}x) = \lim_{n\to\infty}\int_E f_\varepsilon(x)P_n(\mathrm{d}x) = \int_E f_\varepsilon(x)P(\mathrm{d}x)$$

$$\Rightarrow \overline{\lim_{n\to\infty}} P_n(A) = \overline{\lim_{n\to\infty}}\int_E 1_A(x)P_n(\mathrm{d}x) \leqslant \overline{\lim_{n\to\infty}}\int_E f_\varepsilon(x)P_n(\mathrm{d}x)$$

$$= \int_E f_\varepsilon(x)P(\mathrm{d}x) \leqslant P(A_\varepsilon) \downarrow P(A) \ (\text{当 } \varepsilon \downarrow 0 \text{ 时}).$$

即 $\overline{\lim} P_n(A) \leqslant P(A)$ ($\forall$ 可测闭集 A).

现证：(iii)⇒(iv). 事实上，设 $A \in \mathscr{E}$，$\mathring{A} = A \backslash \partial A$，则 $\mathring{A}$ 为开集. 因此，

$$\overline{\lim_{n\to\infty}} P_n([A]) \leqslant P([A]),\ \underline{\lim_{n\to\infty}} P_n(\mathring{A}) \geqslant P(\mathring{A}).$$

若 $P(\partial A) = 0$，则

$$P(\mathring{A}) = P(A) = P([A])$$

$$\Rightarrow \underline{\lim_{n\to\infty}} P_n(A) \geqslant \underline{\lim_{n\to\infty}} P_n(\mathring{A}) \geqslant \overline{\lim_{n\to\infty}} P_n([A]) \geqslant \overline{\lim_{n\to\infty}} P_n(A).$$

故 $P_n \Rightarrow P$.

最后证明：(iv)⇒(i). 事实上，设 f 为 E 上的有界连续函数. 令

$$D = \{y \in \mathrm{R}: P(x \in E: f(x) = y) \neq 0\},$$

则 D 至多为可数集. 的确, 假如 D 无限不可数, 则∃ $\varepsilon>0$, ∍ 集类 $\{f^{-1}(\{y\}): P(f^{-1}(\{y\}))\geqslant\varepsilon,\ y\in D\}$ 无限不可数. 因此, 存在可数集 $\{y_n,\ n\geqslant1\}\subset D$, 使得

$$P(f^{-1}(\{y_n\})) \geqslant \varepsilon\ (\forall\ n \geqslant 1).$$

注意, $\{f^{-1}(\{y\}): y\in D\}$ 在 $\mathscr{E}$ 中是不相联的集类. 于是,

$$\infty = \sum_n P(f^{-1}(\{y_n\})) = P(\bigcup_n f^{-1}(\{y_n\})) \leqslant P(E) = 1.$$

此矛盾表明: D 只能至多可数. 现在构作如下分划: $-M=y_0<y_1<\cdots<y_n=M$, 其中 M 为 f 的界, 且 $y_k \notin D(0\leqslant k\leqslant n)$. 令 $A_k=f^{-1}([y_{k-1},\ y_k))$, 则

$$\partial A_k \subset f^{-1}(\{y_{k-1}\}) \cup f^{-1}(\{y_k\}).$$

从而, $P(\partial A_k)=0\ (y_{k-1},\ y_k\notin D)$. 假如(iv)成立, 则

$$P_m(A_k) \xrightarrow[m\to\infty]{} P(A_k)\ (1 \leqslant k \leqslant n).$$

$$\Rightarrow\left|\int_E f(x)P_m(\mathrm{d}x) - \int_E f(x)P(\mathrm{d}x)\right|$$

$$\leqslant\left|\int_E f(x)P_m(\mathrm{d}x) - \sum_{k=1}^n y_k P_m(A_k)\right|$$

$$+\left|\sum_{k=1}^n y_k P(A_k) - \int_E f(x)P(\mathrm{d}x)\right|$$

$$+\left|\sum_{k=1}^n y_k(P_m(A_k) - P(A_k))\right|$$

$$\leqslant 2\lambda_n + \left|\sum_{k=1}^n y_k(P_m(A_k) - P(A_k))\right|$$

其中 $\lambda_n=\max\{|y_k-y_{k-1}|,\ 1\leqslant k\leqslant n\}$.

$$\Rightarrow \lim_{m\to\infty}\int_E f(x)P_m(\mathrm{d}x) = \int_E f(x)P(\mathrm{d}x), \text{即 } P_n \xrightarrow{w.} P.$$

定理 2.3.2 设 $P,\ P_n,\ n\geqslant1$, 为 $(\mathrm{R},\ \mathscr{B})$ 上的概率测度, 则下列命题等价:

(i) $P_n \xrightarrow{w.} P$；(ii) $P_n \Rightarrow P$；(iii) $F_n \xrightarrow{w.} F$；(iv) $F_n \Rightarrow F$.

证 按定理 2.3.1，仅需证明：

"(ii)：$P_n \Rightarrow P$" 等价于 "(iv)：$F_n \Rightarrow F$."

假定：$P_n \Rightarrow P$. 如果 $P(\{x\})=0$，则 $F_n(x) \to F(x)$. 然而，当 $P(\{x\})=0$ 时，x 为 F 的连续点. 故(iv)成立.

现设 $F_n(x) \to F(x)$ $(\forall\ x \in C(F))$. 如果 A 为开集，则 A 有表示：$A=\bigcup_k I_k$，其中$(I_k,\ k\geqslant 1)$为不相联的开区间列(空集当作开集). 如果 $I_k=(a_k,\ b_k)$非空，即 $a_k<b_k$. 则对$\forall\ \varepsilon>0$，可选子区间 $I'_k=(a'_k,\ b'_k)\subset I_k$，使得

$$P(I_k) \leqslant P(I'_k) + \varepsilon \cdot 2^{-k},\ 且\ a'_k、b'_k \subset C(F).$$

(这是因为 $\mathrm{R}\backslash C(F)$ 至多为可数集). 按 Fatou 引理，

$$\varliminf_{n\to\infty} P_n(A) = \varliminf_{n\to\infty} \sum_k P_n(I_k) \geqslant \sum_k \varliminf_{n\to\infty} P_n(I_k)$$

$$\geqslant \sum_k{}' \varliminf_{n\to\infty} P_n(I'_k),$$

其中 $\sum_k{}'$ 表示对 $I_k \neq \varnothing$ 的指标 k 求和. 按概收敛定义，

$$P_n(I'_k) = F_n(b'_k) - F_n(a'_k) \to F(b'_k) - F(a'_k) = P(I'_k).$$

$\Rightarrow \varliminf_{n\to\infty} P_n(A) \geqslant P(A)$，即 $P_n \Rightarrow P$(按定理 2.3.1).

定义 2.3.3 设$(E,\mathscr{E})$为可测空间. P,Q 为其上的概率测度. $\mathscr{K}_0$ 为 E 上的子集类. 若 $\mathscr{K}_0\subset\mathscr{E}$，且当 $P(A)=Q(A)$ $(\forall\ A\in\mathscr{K}_0)$时，必有 $P=Q$ 于 $\mathscr{E}$ 上，则称 $\mathscr{K}_0$ 为 $\mathscr{E}$ 中的一个**决定类**.

定义 2.3.4 设$(E,\mathscr{E},\rho)$为完备尺度空间. P, P_n, $n\geqslant 1$，均为此空间上的概率测度. 若 E 上的子集类 $\mathscr{K}_1\subset\mathscr{E}$ 具有性质：当 $P_n\Rightarrow P$ 于 $\mathscr{K}_1$ 上时，必有 $P_n\Rightarrow P$ 于 $\mathscr{E}$ 上，则称 $\mathscr{K}_1$ 为 $\mathscr{E}$ 中的一个**收敛决定类**.

例 2.3.1 设$(E,\mathscr{E},\rho)=(\mathrm{R},\mathscr{B},\rho)$，其中 ρ 为常规尺度. 若 $\mathscr{K}_0=\{B:\ B=(-\infty,x],\ x\in\overline{\mathrm{R}}\}$，则按第一章中的定理 1.3.5，$\mathscr{K}_0$ 为 $\mathscr{E}$ 中的一个决定类. 还有 $\sigma(\mathscr{K}_0)=\mathscr{B}$. 按定理 2.3.2，$\mathscr{K}_0$

同样也为$\mathscr{E}$中的一个收敛决定类. 若$\overline{\mathscr{K}}_0=\{B: B=(a,b], a,b\in \overline{\mathrm{R}}\}$, 则同理可知, $\overline{\mathscr{K}}_0$为$\mathscr{E}$中的一个决定类. 同样也有$\sigma(\overline{\mathscr{K}}_0)=\mathscr{B}$.

显然, $\mathscr{B}$本身既为决定类又为收敛决定类. 但讨论$\mathscr{B}$本身的这种性质毫无意义. 之所以引入决定类概念, 原因在于: 要求此类元素少, 结构简单. 类似的方法可以讨论尺度空间$(\mathrm{R}^\infty, \mathscr{B}^\infty, \rho)$中的决定类, 其中尺度$\rho$由下式定义:

$$\rho(x,y)=\sum_{k=1}^{\infty}2^{-k}\rho_1(x_k, y_k),$$

$$\rho_1(x_k, y_k)=\frac{|x_k-y_k|}{1+|x_k-y_k|}$$

$$(x,y\in \mathrm{R}^\infty, x=(x_k, k\geqslant 1), y=(y_k, k\geqslant 1)).$$

令$\mathscr{K}_0=\{\hat{B}: \hat{B}=\mathscr{I}_n(B)$, 其中$B=(-\infty,x], x\in \mathrm{R}^n, n\geqslant 1\}$, 则$\mathscr{K}_0$为$\mathrm{R}^\infty$中的一个决定类, 也是一个收敛决定类.

定理 2.3.3 设$\mathscr{K}_0$为$\mathscr{E}$中的一个决定类, 则

$$\sigma(\mathscr{K}_0)=\mathscr{E}.$$

证 假如结论不真, 则$\exists A_0\in\mathscr{E}$, $\ni A_0\notin\sigma(\mathscr{K}_0)$. 现在, 在$\mathscr{E}$上构作两个概率测度$P$和$Q$, 使得:

$$P(A)=Q(A)\ (\forall A\in\mathscr{K}_0),\ P(A_0)\neq Q(A_0).$$

这种概率测度不难构作. 然而, 按定义 2.3.3, $\mathscr{K}_0$就不是$\mathscr{E}$中的一个决定类, 这与假设矛盾.

定理 2.3.4 在一般尺度空间中柱集类可以是决定类而不是收敛决定类.

证 下面通过构作一反例来证实这一结论. 考虑尺度空间$(C, \mathscr{B}_0(C), \rho)$, 其中$C\subset \mathrm{R}^T$ $(T\doteq[0,1])$为连续函数类. ρ为尺度:

$$\rho(x,y)=\sup_{t\in T}|x(t)-y(t)|\ (x,y\in C).$$

$\mathscr{B}_0(C)$为关于尺度ρ的开集类生存的σ-代数. 前一章中已证明: $\mathscr{B}(C)=\mathscr{B}_0(C)$, 其中$\mathscr{B}(C)$为由柱集类生成的$\sigma$-代数. 设$P$为集

中在 $x_0=0$ 处的概率测度，即

$$P(B)=\delta(x_0,B)=\begin{cases}1, & x_0\in B,\\ 0, & x_0\notin B.\end{cases}\quad (B\in\mathscr{B}_0(C)).$$

P_n 为集中在 x_n 处的概率测度，其中

$$x_n=x_n(t)=nt1_{\left[0,\frac{1}{n}\right]}(t)+(2-nt)1_{\left(\frac{1}{n},\frac{2}{n}\right)}(t).$$

假如 A 为柱集，且 $P(\partial A)=0$，则 $P_n(A)\to P(A)$. 事实上，若 $x_0=0\in A$，则 $x_0\in \mathring{A}$（A 的内点集），且 $P(\mathring{A})=1$. 注意，A 为柱集. 因此，当 n 足够大时，必有 $x_n\in\mathring{A}$. 从而，$P_n(\mathring{A})=1$. 于是，$P_n(A)\to P(A)$. 若 $x_0\notin A$，则 $x_0\notin[A]$（闭包），且当 n 足够大时，$x_n\notin[A]$. 由此知：$P_n(A)\to 0=P(A)$. 这就证实 $P_n\Rightarrow P$ 于 $\mathscr{K}_0$ 上，其中 $\mathscr{K}_0$ 为 C 上的柱集类.

现在考虑 $\mathscr{B}_0(C)$ 中的下形元素：

$$A'=\left\{\alpha\in C:\ |\alpha(t)|\leqslant\frac{1}{2}\ (\forall\ t\in T)\right\}.$$

则 $x_0=0\in A'$. 从而，$P(A')=1$，$P(\partial A')=0$. 注意，对$\forall\ n\geqslant 1$，有 $x_n\notin A'$. 因此，$P_n(A')=0$ $(\forall\ n\geqslant 1)$. 故

$$P_n(A')\not\to P(A').$$

从而，$P_n\not\Rightarrow P$ 于 $\mathscr{B}_0(C)$ 上.

综上分析，$\mathscr{K}_0$ 为 $\mathscr{B}_0(C)$ 中的一个决定类. 但不是收敛决定类.

注 4 此例表明：$\mathscr{K}_0\subset\mathscr{K}_1$ 可以是真包含.

§2.4 概率测度的相对紧性与紧箍性

假定 $\mathscr{P}=(P_n,\ n\geqslant 1)$ 为常规尺度空间 $(\mathrm{R},\mathscr{B},\rho)$ 上的概率测度族，且 P_n 集中在点 $\{n\}\in\mathscr{B}$ 处，则对$\forall\ a,b\in\mathrm{R}$，有 $P_n(a,b]\xrightarrow[n\to\infty]{}0$，$P_n(\mathrm{R})=1\longrightarrow 1$，

$$F_n(x)=P_n(-\infty,x]\xrightarrow[n\to\infty]{}0\ (\forall\ x\in\mathrm{R}).$$

这个事实表明：$\mathscr{P}$ 中不存在弱收敛子列. 按定理 2.3.2，R 上的分布函数空间在弱收敛意义下不是一个紧空间. 这就提出了一个问题，在什么条件下，分布函数弱收敛到某个分布函数.

本节总假定$(E,\mathscr{E},\rho)$为一个完备尺度空间.

定义 2.4.1 设$\mathscr{P}=(P_\alpha,\ \alpha\in\mathscr{U})$为$(E,\mathscr{E},\rho)$上的概率测度族. 若 $\mathscr{P}$ 中任一无穷子列都含有一子列弱收敛到某个概率测度，则称 $\mathscr{P}$ **是相对紧**的（"相对"一词表示：弱收敛结果不一定在 $\mathscr{P}$ 中）.

若对$\forall\ \varepsilon>0$，$\exists$ 紧集$K\in\mathscr{E}$，使得

$$\sup\{P_\alpha(E\backslash K),\ \alpha\in\mathscr{U}\}<\varepsilon,$$

则称 $\mathscr{P}$ **是紧箍的**.

若 $\mathscr{P}$ 相对紧，则称相应的分布族$\mathscr{F}=(F_\alpha,\ \alpha\in\mathscr{U})$亦是相对紧的.

若 R 上取值于$[0,1]$的实函数 G 具有下列性质：右连左极，且$0\leqslant G(-\infty)\leqslant G(+\infty)\leqslant 1$，其中 $G(\pm\infty)=\lim\limits_{x\to\pm\infty}G(x)$，则称 G 为 R 上的**广义分布函数**.

定理 2.4.1（Helly 定理） 设 $\mathscr{G}$ 为 R 上的广义分布函数全体，则 $\mathscr{G}$ 在概收敛意义下是序列紧的.

证 $\mathscr{G}$ 序列紧即是对 $\mathscr{G}$ 中任一无穷子列$(G_n,\ n\geqslant 1)$，$\exists$ 子列$(G_{n_k},\ k\geqslant 1)$，$\ni$ $G_{n_k}\Rightarrow G\in\mathscr{G}$，以下分二步证明这一事实.

第一步构作子列，设$T=(x_n,\ n\geqslant 1)$为 R 中的一个可数稠集. 选子列$(G_{n_k^1},\ k\geqslant 1)$，$\ni$ $\lim\limits_{k\to\infty}G_{n_k^1}(x_1)=g_1$. 接着在$(G_{n_k^1})$中选子列$(G_{n_k^2},\ k\geqslant 1)$，$\ni$ $\lim\limits_{k\to\infty}G_{n_k^2}(x_2)=g_2$. 继续这一作法，可选出一串子列$(G_{n_k^m},\ k\geqslant 1)$ $(m\geqslant 1)$，具有如下性质：

$$(G_{n_k^{m+1}},\ k\geqslant 1)\subset(G_{n_k^m},\ k\geqslant 1),\ \lim_{k\to\infty}G_{n_k^m}(x_m)=g_m.$$

按对角线法，可选子列$(G_{n_m^m},\ m\geqslant 1)$，使得

$$\lim_{m\to\infty}G_{n_m^m}(x_l)=g_l(\forall\ l\geqslant 1).$$

现在，设 G_T 为 R 上的实函数，且 $G_T(x_l)=g_l(\forall\ x_l\in T)$.

令 $$G(x)=\inf\{G_T(y):\ x<y\in T\}\ (x\in\mathbf{R}).$$

显然,$G\in\mathscr{G}$.

第二步证明:$G_{n_m^m}\Rightarrow G$(概收敛). 事实上,让 $C(G)$ 为 G 的连续点全体,$x^0\in C(G)$. 若 $x^0<y\in T$,则

$$\overline{\lim_{m\to\infty}}G_{n_m^m}(x^0)\leqslant\overline{\lim_{m\to\infty}}G_{n_m^m}(y)=G_T(y)$$

$$\Rightarrow\overline{\lim_{m\to\infty}}G_{n_m^m}(x^0)\leqslant\inf\{G_T(y):\ x^0<y\in T\}=G(x^0).$$

另一方面,对$\forall\ x^1<x^0$,有

$$G(x^1)\leqslant\underline{\lim_{m\to\infty}}G_{n_m^m}(x^0)$$

$$\Rightarrow G(x^0-)\leqslant\underline{\lim_{m\to\infty}}G_{n_m^m}(x^0)\leqslant\overline{\lim_{m\to\infty}}G_{n_m^m}(x^0)\leqslant G(x^0)$$

$$\Rightarrow\lim_{m\to\infty}G_{n_m^m}(x^0)=G(x^0)\ (\forall\ x^0\in C(G)).$$

定理 2.4.2(Prohorov 定理). 设 $\mathscr{P}=(P_\alpha,\ \alpha\in\mathscr{U})$是定义在完备可分尺度空间$(E,\mathscr{E},\rho)$上的概率测度族,则“$\mathscr{P}$ 相对紧”等价于“$\mathscr{P}$ 紧箍”.

证 为简明计,不妨设$(E,\mathscr{E},\rho)=(\mathbf{R},\mathscr{B},\rho)$.

首先假定 $\mathscr{P}$ 相对紧,采用反证法. 假如 $\mathscr{P}$ 不紧箍,则$\exists\ \varepsilon>0$,$\ni$ 对任意紧集 $K\in\mathscr{B}$,有

$$\sup\{P_\alpha(\mathbf{R}\backslash K),\ \alpha\in\mathscr{U}\}>\varepsilon.$$

令 $K_n=[-n,n]$,则 K_n 为紧集. 从而,对$\forall\ n\geqslant1$,有

$$\sup\{P_\alpha(\mathbf{R}\backslash K_n),\ \alpha\in\mathscr{U}\}>\varepsilon.$$

令 $\mathring{K}_n=(-n,n)$,则更有

$$\sup\{P_\alpha(\mathbf{R}\backslash\mathring{K}_n),\ \alpha\in\mathscr{U}\}>\varepsilon.\ (\forall\ n\geqslant1)$$

于是,$\exists\ (\alpha_n,\ n\geqslant1)\subset\mathscr{U}$, $\ni P_{\alpha_n}(\mathbf{R}\backslash\mathring{K}_n)>\varepsilon\ (\forall\ n\geqslant1)$. 由 $\mathscr{P}$ 相对紧知:$\exists\ (\alpha_{n_k},\ k\geqslant1)\subset(\alpha_n,\ n\geqslant1)$, $\ni P_{\alpha_{n_k}}\xrightarrow{w.}Q$. 注意,$\mathbf{R}\backslash\mathring{K}_n$ 为闭集. 按定理 2.3.1,有

$$\varepsilon\leqslant\overline{\lim_{k\to\infty}}P_{\alpha_{n_k}}(\mathbf{R}\backslash\mathring{K}_n)\leqslant Q(\mathbf{R}\backslash\mathring{K}_n)\ (\forall\ n\geqslant1).$$

然而,$Q(\mathrm{R}\backslash\mathring{K}_n)\downarrow 0$ 当 $n\uparrow\infty$ 时,这就出现矛盾.

现在假定 $\mathscr{P}$ 紧箍,$\mathscr{F}=(F_\alpha,\ \alpha\in\mathscr{U})$ 为相应于 $\mathscr{P}$ 的分布函数族. $(P_n,\ n\geqslant 1)\subset\mathscr{P}$ 相应于 $(F_n,\ n\geqslant 1)\subset\mathscr{F}$. 按 Helly 定理,存在子列 $(F_{n_k},\ k\geqslant 1)$ 概收敛于广义分布函数 G,即 $F_{n_k}(x)\xrightarrow[k\to\infty]{}G(x)$ $(\forall\ x\in C(G))$. 若 G 为分布函数,则按定理 2.3.2,$P_{n_k}\xrightarrow{w.}Q$,其中 Q 是由 G 产生的概率测度. 这表明:$\mathscr{P}$ 是相对紧的. 现在剩下的问题是证明:G 为分布函数. 为此,仅需证明:$G(-\infty)=0$, $G(+\infty)=1$. $\mathscr{P}$ 紧箍意味着:对 $\forall\ \varepsilon>0$,$\exists\ a,b\in\mathrm{R}$,使得

$$
\begin{aligned}
&\sup\{P_\alpha(\mathrm{R}\backslash[a,b]),\ \alpha\in\mathscr{U}\}<\varepsilon\\
\Rightarrow&\sup\{P_\alpha(\mathrm{R}\backslash(\bar{a},b]),\ \alpha\in\mathscr{U}\}<\varepsilon\ (\forall\ \bar{a}<a).
\end{aligned}
$$

让 $a',b'\in C(G)$,且 $a'\leqslant\bar{a}$, $b\leqslant b'$,则

$$
\begin{aligned}
1-\varepsilon&<1-P_{n_k}(\mathrm{R}\backslash(\bar{a},b])\\
&=F_{n_k}(b)-F_{n_k}(\bar{a})\leqslant F_{n_k}(b')-F_{n_k}(a')\\
\Rightarrow&1-\varepsilon\leqslant G(b')-G(a')\\
\Rightarrow&1-\varepsilon\leqslant G(\infty)-G(-\infty)\\
\Rightarrow&G(\infty)-G(-\infty)=1.
\end{aligned}
$$

故按 $G\in\mathscr{G}$ 知:$G(-\infty)=0$, $G(\infty)=1$.

定理 2.4.3 设 $P_\alpha=N(m_\alpha,\sigma_\alpha^2)$, $\alpha\in\mathscr{U}$, 为 $(\mathrm{R},\mathscr{B},\rho)$ 上的 Gauss 概率测度族. 则 $\mathscr{P}$ 紧箍的充要条件是:

$\exists$ 有限的 $a>0$, $b>0$, $\ni\ \sup\limits_\alpha|m_\alpha|\leqslant a$, $\sup\limits_\alpha\sigma_\alpha^2\leqslant b$.

证 令 $K_n=[-n,n]$,则

$$
\begin{aligned}
P_\alpha(\mathrm{R}\backslash K_n)=&\frac{1}{\sqrt{2\pi}\sigma_\alpha}\Big[\int_n^\infty\exp\Big(-\frac{(x-m_\alpha)^2}{2\sigma_\alpha^2}\Big)\mathrm{d}x\\
&+\int_{-\infty}^{-n}\exp\Big(-\frac{(x-m_\alpha)^2}{2\sigma_\alpha^2}\Big)\mathrm{d}x\Big]\\
\leqslant&\sqrt{\frac{2}{\pi}}\int_n^\infty\exp\Big(-\frac{(x-|m_\alpha|)^2}{2\sigma_\alpha^2}\Big)\mathrm{d}x\cdot\frac{1}{\sigma_\alpha}
\end{aligned}
$$

$$=\sqrt{\frac{2}{\pi}}\int_{t'_a}^{\infty}\exp\left(-\frac{x^2}{2}\right)\mathrm{d}x,\ \text{其中}\ t'_a=\frac{n-|m_a|}{\sigma_a}.$$

$$P_a(\mathrm{R}\backslash K_n)\geqslant\frac{1}{\sqrt{2\pi}}\int_{t_a^0}^{\infty}\exp\left(-\frac{x^2}{2}\right)\mathrm{d}x,\ \text{其中}\ t_a^0=\frac{n-m_a}{\sigma_a}.$$

由上两式即可得所要求的结论.

§2.5 弱收敛与特征函数

引理2.5.1 设定义在$(\mathrm{R},\mathscr{B},\rho)$上的概率测度族$\mathscr{P}=(P_n,\ n\geqslant1)$是紧箍的. 若$\mathscr{P}$的每个弱收敛子列$(P_{n_k},\ k\geqslant1)$都收敛到同一概率测度$P$,则全序列$\mathscr{P}$弱收敛到$P$.

证 设$P_n \overset{w.}{\not\to} P$. 则$\exists$ 一个有界连续函数f_0,使得

$$\int_{\mathrm{R}}f_0(x)P_n(\mathrm{d}x)\not\to\int_{\mathrm{R}}f_0(x)P(\mathrm{d}x),$$

即$\exists\ \varepsilon>0$及无穷子列$(n_k,\ k\geqslant1)$,使得

$$\left|\int_{\mathrm{R}}f_0(x)P_{n_k}(\mathrm{d}x)-\int_{\mathrm{R}}f_0(x)P(\mathrm{d}x)\right|>\varepsilon.$$

然而,按定理2.4.2,$\exists$ 子列$(a_k,\ k\geqslant1)\subset(n_k,\ k\geqslant1)$,使得

$$P_{a_k}\xrightarrow{w.}Q.$$

按假设$Q=P$,因此

$$\int_{\mathrm{R}}f_0(x)P_{a_k}(\mathrm{d}x)\longrightarrow\int_{\mathrm{R}}f_0(x)P(\mathrm{d}x).$$

这与子列$(P_{n_k},\ k\geqslant1)$的性质矛盾.

引理2.5.2 设定义在$(\mathrm{R},\mathscr{B},\rho)$上的概率测度族$\mathscr{P}=(P_n,\ n\geqslant1)$是紧箍的. $\Phi=(\varphi_n,\ n\geqslant1)$是相应于$\mathscr{P}$的特征函数族,$\varphi_n(t)=\int_{\mathrm{R}}\exp(itx)P_n(\mathrm{d}x)\ (\forall\ t\in\mathrm{R},\ n\geqslant1)$.

则“$P_n\xrightarrow{w.}P$”等价于“$\lim\limits_{n\to\infty}\varphi_n(t)$ 存在 $(\forall\ t\in\mathrm{R})$”.

证 按弱收敛之定义，由 $P_n \xrightarrow{w.} P$ 立即可推出$\lim\limits_{n\to\infty}\varphi_n(t)$存在$(\forall\ t\in \mathrm{R})$. 反之，假定对$\forall\ t\in\mathrm{R}$，有$\lim\limits_{n\to\infty}\varphi_n(t)$存在. 如若 $P_n \not\xrightarrow{w.} P$. 则按定理 2.4.2 和引理 2.5.1，∃ 子列$(P_{n_k},\ k\geqslant 1)$弱收敛到 P，子列$(P_{a_k},\ k\geqslant 1)$弱收敛到 Q，→ $P\neq Q$. 按条件，应有

$$\int_{\mathrm{R}}\exp(itx)P(\mathrm{d}x) = \int_{\mathrm{R}}\exp(itx)Q(\mathrm{d}x) \quad (\forall\ t\in\mathrm{R}).$$

按特征函数的唯一性定理，有 $P=Q$. 这就出现矛盾. 故 $P_n \xrightarrow{w.} P$.

引理 2.5.3 设 F 为 R 上的一个分布函数. φ 为其特征函数，则∃ $K>0$，使得

$$\int_{\{|x|>\frac{1}{a}\}}\mathrm{d}F(x) \leqslant \frac{K}{a}\int_0^a(1-\mathrm{Re}\varphi(t))\mathrm{d}t \quad (\forall\ a>0).$$

证 易证

$$\begin{aligned}
&\inf_{|y|\geqslant 1}\left(1-\frac{\sin y}{y}\right) = 1-\sin 1\\
\Rightarrow &\frac{1}{a}\int_0^a(1-\mathrm{Re}\varphi(t))\mathrm{d}t\\
&= \frac{1}{a}\int_0^a\int_{\mathrm{R}}(1-\cos(tx))\mathrm{d}F(x)\mathrm{d}t\\
&= \int_{\mathrm{R}}\left(\frac{1}{a}\int_0^a(1-\cos(tx))\mathrm{d}t\right)\mathrm{d}F(x)\\
&= \int_{\mathrm{R}}\left(1-\frac{\sin(ax)}{ax}\right)\mathrm{d}F(x)\\
&\geqslant \inf_{|y|\geqslant 1}\left(1-\frac{\sin y}{y}\right)\int_{\{|x|>\frac{1}{a}\}}\mathrm{d}F(x)\\
&= \frac{1}{K}\int_{\{|x|\geqslant\frac{1}{a}\}}\mathrm{d}F(x)，其中\ K=(1-\sin 1)^{-1}.
\end{aligned}$$

定理 2.5.1（连续性定理） 设$\mathscr{F}=(F_n,\ n\geqslant 1)$为 R 上的分布函数族. $\Phi=(\varphi_n,\ n\geqslant 1)$为相应于$\mathscr{F}$ 的特征函数族.

(i) 若 $F_n\Rightarrow F$，其中 F 为分布函数. 则

$$\varphi_n(t) \to \varphi(t) \quad (\forall\, t \in \mathrm{R}),$$

其中 φ 为 F 的特征函数.

(ii) 若 $\varphi(t)=\lim\limits_{n\to\infty}\varphi_n(t)$存在($\forall\, t\in\mathrm{R}$),且 $\varphi(t)$连续于 $t=0$,则 φ 为某分布函数 F 的特征函数,且 $F_n\Rightarrow F$.

证 结论(i)可直接由弱收敛之定义得出. 下证结论(ii). 设 $\mathscr{P}=(P_n,\ n\geqslant 1)$是相应于 $\mathscr{F}$ 的概率测度族. 若 $\mathscr{P}$ 紧箍,且 $\lim\limits_{n\to\infty}\varphi_n(t)$存在($\forall\, t\in\mathrm{R}$). 则按引理 2.5.2,$P_n\xrightarrow{w.}P$,这里 P 为概率测度. 按结论(i),$\varphi_n(t)\to\varphi(t)$ ($\forall\, t\in\mathrm{R}$),其中 φ 为 P 的特征函数. 因此,剩下的问题是证明:$\mathscr{P}$ 紧箍.

令 $\mathring{B}_a=\left(-\dfrac{1}{a},\dfrac{1}{a}\right)$. 按引理 2.5.3,有

$$P_n(\mathrm{R}\backslash\mathring{B}_a)=\int_{\{|x|\geqslant\frac{1}{a}\}}\mathrm{d}F(x)\leqslant\frac{k}{a}\int_0^a(1-\mathrm{Re}\varphi_n(t))\mathrm{d}t\ (\forall\ a>0).$$

按 Lebesque 控制收敛定理,有

$$\lim_{n\to\infty}\frac{k}{a}\int_0^a(1-\mathrm{Re}\varphi_n(t))\mathrm{d}t=\frac{k}{a}\int_0^a(1-\mathrm{Re}\varphi(t))\mathrm{d}t\quad(\forall\ a>0).$$

按假设 φ 连续于 $t=0$. 因此,对$\forall\ \varepsilon>0$,$\exists\ a_0=a(\varepsilon)>0$,使得

$$|1-\mathrm{Re}\varphi(t)|\leqslant\frac{1}{k}\varepsilon\ (\forall\ 0\leqslant t\leqslant a\leqslant a_0).$$

又对$\forall\ a>0$,存在 $N=N(\varepsilon,a)\geqslant 1$,使得

$$P_n(\mathrm{R}\backslash\mathring{B}_a)\leqslant\frac{k}{a}\int_0^a(1-\mathrm{Re}\varphi(t))\mathrm{d}t+\varepsilon\ (\forall\ n\geqslant N).$$

综上分析,得

$$P_n(\mathrm{R}\backslash[\mathring{B}_a])\leqslant P_n(\mathrm{R}\backslash\mathring{B}_a)\leqslant 2\varepsilon\ (0<a,\ n\geqslant N).$$

$\Rightarrow \mathscr{P}$ 紧箍.

定理 2.5.2(大数定律) 设 $\xi=(\xi_n,\ n\geqslant 1)$独立同分布,且 $E|\xi_1|<\infty$. 令 $m=E\xi_1$, $S_n=\sum\limits_{k=1}^n\xi_k$,则$\dfrac{1}{n}S_n\xrightarrow{P.}m$.

证 令 $\varphi(t)=E\exp(it\xi_1)$ ($t\in\mathrm{R}$),

$$\eta_n = \frac{1}{n}S_n,\ \varphi_n(t) = E\exp(it\eta_n)\ (t \in \mathrm{R}).$$

按 ξ 独立同分布，$\varphi_n(t)=\left(\varphi\left(\frac{t}{n}\right)\right)^n$. 显然

$$\varphi(t) = 1 + itm + 0(t)\ (\text{当 } |t| \text{ 足够小时}).$$

因此，对于 $|t|\leqslant N<\infty$，当 n 足够大时，有

$$\varphi\left(\frac{t}{n}\right) = 1 + i\frac{t}{n}m + 0\left(\frac{t}{n}\right)$$

$$\Rightarrow \varphi_n(t) = \left(1 + i\frac{t}{n}m + 0\left(\frac{t}{n}\right)\right)^n \xrightarrow[n\to\infty]{} \exp(itm)$$

$$= \varphi^*(t)\ (\forall\ |t| \leqslant N < \infty).$$

即
$$\varphi_n(t) \xrightarrow[n\to\infty]{} \varphi(t)\ (\forall\ t \in \mathrm{R}).$$

显然，$\varphi^*(t)$连续于 $t=0$，且 $\varphi^*(0)=1$. 按定理 2.5.1(ii)，φ^* 为特征函数，相应的概率测度为

$$P^*(B) = 1_B(m)\quad (\forall\ B \in \mathscr{B}),$$

且 $P_n \xrightarrow{w.} P$，即 $F_n \xrightarrow{w.} F$. 定义函数 f_ε:

$$f_\varepsilon(x) = 1_{(-\varepsilon,\varepsilon)}(x-m) + \left(2 - \frac{|x-m|}{\varepsilon}\right)1_{[\varepsilon,2\varepsilon)}(|x-m|),$$

则对$\forall\ \varepsilon>0$，f_ε 为有界连续函数，且

$$\int_{\mathrm{R}} f_\varepsilon(x)P^*(\mathrm{d}x) = 1 \Rightarrow \int_{\mathrm{R}} f_\varepsilon(x)P_n(\mathrm{d}x) \xrightarrow[n\to\infty]{} 1$$

$$\Rightarrow 1 = \varliminf_{n\to\infty}\int_{\mathrm{R}} f_\varepsilon(x)P_n(\mathrm{d}x) \leqslant \varliminf_{n\to\infty} P_n(|x-m| < 2\varepsilon) = 1$$

$$\Rightarrow \lim_{n\to\infty} P_n(|x-m| < 2\varepsilon) = 1,\text{即}\frac{1}{n}S_n \xrightarrow{P.} m.$$

定理 2.5.3(中心极限定理)　设 $\xi=(\xi_n,\ n\geqslant1)$独立同分布，$m=E\xi_1$，$\sigma^2=E(\xi_1-E\xi_1)^2<\infty$，$S_n=\sum_{k=1}^n \xi_k$，则

$$\frac{S_n - nm}{\sigma\sqrt{n}} \xrightarrow{w.} N,\text{其中 } N \text{ 服从正态分布 } N(0,1).$$

证　不妨设 $m=0$. 让

$$\varphi_n(t)=E\exp\left(it\frac{S_n}{\sigma\sqrt{n}}\right),\quad \varphi\left(\frac{t}{\sigma\sqrt{n}}\right)=E\exp\left(i\frac{t}{\sigma\sqrt{n}}\xi_1\right).$$

则 $\varphi_n(t)=\varphi^n\left(\frac{t}{\sigma\sqrt{n}}\right)$. 易证

$$\varphi\left(\frac{t}{\sigma\sqrt{n}}\right)=1-\frac{t^2}{2n}+\beta\left(\frac{t}{\sigma\sqrt{n}}\right),\text{其中}\frac{\beta(t)}{t^2}\xrightarrow[t\to 0]{}0.$$

设$(z_k,\ 1\leqslant k\leqslant n)$和$(w_k,\ 1\leqslant k\leqslant n)$为两族其模不超过 1 的复数,则

$$\begin{aligned}\left|\prod_{k=1}^n z_k-\prod_{k=1}^n w_k\right|&=\left|(z_1-w_1)\prod_{k=2}^n z_k+w_1\left(\prod_{k=2}^n z_k-\prod_{k=2}^n w_k\right)\right|\\&\leqslant|z_1-w_1|+\left|\prod_{k=2}^n z_k-\prod_{k=2}^n w_k\right|\\&\leqslant\cdots\leqslant\sum_{k=1}^n|z_k-w_k|\\&\Rightarrow\left|\varphi^n\left(\frac{t}{\sigma\sqrt{n}}\right)-\left(1-\frac{t^2}{2n}\right)^n\right|\\&\leqslant n\left|\varphi\left(\frac{t}{\sigma\sqrt{n}}\right)-\left(1-\frac{t^2}{2n}\right)\right|\quad\left(\frac{t^2}{2n}<1\right).\end{aligned}$$

由$\frac{1}{t^2}\beta(t)\xrightarrow[t\to 0]{}0$ 可知

$$n\left|\varphi\left(\frac{t}{\sigma\sqrt{n}}\right)-\left(1-\frac{t^2}{2n}\right)\right|\xrightarrow[n\to\infty]{}0\quad(\forall\, t\in\mathrm{R})$$

$$\Rightarrow\lim_{n\to\infty}\varphi_n(t)=\lim_{n\to\infty}\left(1-\frac{t^2}{2n}\right)^n=\mathrm{e}^{-\frac{t^2}{2}}\quad(\forall\, t\in\mathrm{R}).$$

按定理 2.5.1(ii),便可得所要求的结论.

§2.6 大偏差与重对数律

大偏差研究变量离开均值很大时的准确上下界问题,目的在于利用这种界探讨大数定律中的收敛速率,并由此导致重对数律.

（一）矩生成函数

设 X 为定义在$(\Omega,\mathscr{F},P)$上的 r.v. 令

$$M(t) = E\exp(tX) \quad (\forall\, t \in \mathrm{R}),$$
$$C(t) = \log M(t) \quad (\forall\, t \in \mathrm{R}).$$

则分别称 $M(t)$和 $C(t)$为 X 的**矩生成函数**和**累积生成函数**. 显然有

$$EX^k = M^{(k)}(0);\ C'(t) = \frac{M'(t)}{M(t)},\ C''(t) = \frac{M''(t)}{M(t)} - \frac{(M'(t))^2}{(M(t))^2},$$
$$C(0) = 0,\quad C'(0) = EX,\quad C''(0) = EX^2 - (EX)^2.$$

此外,还具有下列有用的性质:

(i) $M(t)$唯一决定 X 的概率分布 P_X;

(ii) $M(t)$和 $C(t)$均为下凸函数.

事实上,$M''(t)=EX^2e^{tX}\geqslant 0$;

$$C''(t) = \frac{1}{(M(t))^2}(M''(t)M(t) - (M'(t))^2);$$
$$\begin{aligned} M''(t)M(t) &= \int_{\mathrm{R}^2} x_1^2 e^{t(x_1+x_2)} P_X(dx_1)P_X(dx_2) \\ &= \frac{1}{2}\int_{\mathrm{R}^2} (x_1^2 + x_2^2) e^{t(x_1+x_2)} P_X(dx_1)P_X(dx_2) \\ &\geqslant \int_{\mathrm{R}^2} x_1x_2 e^{t(x_1+x_2)} P_X(dx_1)P_X(dx_2) = (M'(t))^2 \end{aligned}$$

$\Rightarrow C''(t)\geqslant 0$. 从而 $M(t)$和 $C(t)$为下凸函数.

（二）大偏差

设 Y 为具有 0 均值的 r.v. 大偏差问题就在于估计概率 $P(Y\geqslant\alpha)$ $(\alpha>0)$. 这等价于:"Y 具有均值$-\alpha(\alpha>0)$,估计概率 $P(Y\geqslant 0)$."

定理 2.6.1 设 $EY<0$, $P(Y>0)>0$. 则

$$\rho \geqslant P(Y \geqslant 0) \geqslant \rho P(Z \geqslant 0) \exp\left(- \frac{\tau S}{P(Z \geqslant 0)}\right),$$

其中 τ 由等式 $\rho=M(\tau)=\inf_t M(t)$ 决定.

$$0<\rho<1,\quad 0<\tau<\infty,$$

r.v. Z 具有如下概率分布:

$$P_Z(\mathrm{d}y)=\frac{1}{\rho}\mathrm{e}^{\tau y}P_Y(\mathrm{d}y),\quad S=(EZ^2)^{\frac{1}{2}}.$$

证 首先由等式 $M(\tau)=\inf_t M(t)$ 定义 τ. $\rho=M(\tau)$, 则 $0<\rho<1$, 且 $\tau>0$ 为有限实数. 事实上, 注意,

$$M(0)=1,\ M(t)\text{ 下凸},\ M'(0)=EY<0.$$

则 $M(t)$ 在 0 附近单调下降. 于是

$$\inf_t M(t)<1,\quad \tau>0.$$

由 $P(Y>0)>0$ 知: $\exists\ y_0>0$, 使 $P(Y\geqslant y_0)>0$. 于是

$$M(t)\geqslant E(\mathrm{e}^{tY}1_{\{Y\geqslant y_0\}})\geqslant \mathrm{e}^{ty_0}P(Y\geqslant y_0)\Rightarrow M(t)\xrightarrow[t\to\infty]{}\infty.$$

这表明: τ 为有限实数. 显然, 对 $\forall\ t\in\mathrm{R}$, 有 $M(t)>0$. 于是 $\rho>0$. 故 $0<\rho<1$.

现在定义辅助变量 Z, 使得

$$P_Z(B)=\frac{1}{\rho}\int_B \mathrm{e}^{\tau y}P_Y(\mathrm{d}y)\quad(\forall\ B\in\mathscr{B}),$$

其中 P_Z, P_Y 分别为 Z, Y 的概率分布. 则 Z 具有下列特点:

$$EZ=\int_{\mathrm{R}} zP_Z(\mathrm{d}z)=\int_{\mathrm{R}} y\frac{1}{\rho}\mathrm{e}^{\tau y}P_Y(\mathrm{d}y)=\frac{1}{\rho}M'(\tau)=0;$$

$$M_Z(t)\triangleq E\mathrm{e}^{tZ}=\int_{\mathrm{R}}\mathrm{e}^{ty}\frac{1}{\rho}\mathrm{e}^{\tau y}P_Y(\mathrm{d}y)=\frac{1}{\rho}M(\tau+t);$$

$$S^2=EZ^2=\int_{\mathrm{R}}y^2\frac{1}{\rho}\mathrm{e}^{\tau y}P_Y(\mathrm{d}y)=\frac{1}{\rho}M''(\tau)>0.$$

我们证明: $P(Y\geqslant 0)\leqslant\rho$. 事实上

$$P(Y\geqslant 0)=P(\mathrm{e}^{tY}\geqslant 1)\leqslant E\mathrm{e}^{tY}=M(t)\quad(\forall\ t>0).$$

让 $t=\tau$, 由此式便可得 $P(Y\geqslant 0)\leqslant\rho$.

最后讨论 $P(Y\geqslant 0)$ 的向下不等式. 令 $p=P(Z\geqslant 0)$. 注意, $P_Y(\mathrm{d}y)=\rho\mathrm{e}^{-\tau y}P_Z(\mathrm{d}y)$, 则可得

$$P(Y \geqslant 0) = \int_{\{y \geqslant 0\}} P_Y(\mathrm{d}y) = \int_{\{y \geqslant 0\}} \rho \mathrm{e}^{-\tau y} P_Z(\mathrm{d}y) = \rho \mathrm{e}^{-\theta},$$

其中 $\mathrm{e}^{-\theta} = E(\mathrm{e}^{-\tau Z} 1_{\{Z \geqslant 0\}})$. 注意,$\log x$ 为上凸函数,$-\log x$ 为下凸函数. 定义辅变量 $\widetilde{Z}$,使得

$$P_{\widetilde{Z}}(B) = \frac{1}{p} P_Z(B \cap \{Z \geqslant 0\}) \quad (P_{\widetilde{Z}} \text{ 为 } \widetilde{Z} \text{ 的概率分布}),$$

则
$$\mathrm{e}^{-\theta} = \int_{\mathbf{R}} \mathrm{e}^{-\tau x} p P_{\widetilde{Z}}(\mathrm{d}x) = p E \mathrm{e}^{-\tau \widetilde{Z}}.$$

按 Jensen 不等式,有

$$E(-\log \mathrm{e}^{-\tau \widetilde{Z}}) \geqslant -\log(E\mathrm{e}^{-\tau \widetilde{Z}})$$

$$\Rightarrow \log(E\mathrm{e}^{-\tau \widetilde{Z}}) \geqslant -E(\tau \widetilde{Z}) = -\frac{\tau}{p} \int_{\{y \geqslant 0\}} y P_Z(\mathrm{d}y)$$

$$= -\frac{\tau S}{p} \int_{\{y \geqslant 0\}} \frac{y}{S} P_Z(\mathrm{d}y) \geqslant -\frac{\tau S}{p}$$

$$\Rightarrow -\theta = \log p + \log(E\mathrm{e}^{-\tau \widetilde{Z}}) \geqslant \log p - \frac{\tau S}{p}$$

$$= -\frac{\tau S}{P(Z \geqslant 0)} + \log P(Z \geqslant 0)$$

$$\Rightarrow \theta \leqslant \frac{\tau S}{P(Z \geqslant 0)} - \log P(Z \geqslant 0).$$

定理 2.6.2(估计 $P(Z \geqslant 0)$ 的下界) 设 $EZ=0$, $S^2=EZ^2$, $\xi^4=EZ^4$,则 $p=P(Z \geqslant 0) \geqslant \frac{S^4}{4\xi^4}$.

证 设 $Z^+ = Z \vee 0$. $Z^- = Z^+ - Z$. 则 $Z^\pm$ 为非负 r.v. 利用 Hölder 不等式,可得下列各式:

$$S^2 = EZ^2 = E(Z^+)^2 + E(Z^-)^2;$$

$$E(Z^+)^2 \leqslant (E1_{\{Z \geqslant 0\}})^{\frac{1}{2}} (EZ^4)^{\frac{1}{2}} = p^{\frac{1}{2}} \xi^2;$$

$$E(Z^-)^2 = E(Z^-)^{\frac{2}{3}} (Z^-)^{\frac{4}{3}} \leqslant (EZ^-)^{\frac{2}{3}} (E(Z^-)^4)^{\frac{1}{3}} \leqslant (EZ^-)^{\frac{2}{3}} \xi^{\frac{4}{3}};$$

$$EZ^- = EZ^+ \leqslant p^{\frac{3}{4}} (EZ^4)^{\frac{1}{4}} = \xi \cdot p^{\frac{3}{4}}$$

$$\Rightarrow S^2 \leqslant p^{\frac{1}{2}} \xi^2 + (EZ^-)^{\frac{2}{3}} \xi^{\frac{4}{3}} \leqslant p^{\frac{1}{2}} \xi^2 + p^{\frac{1}{2}} \xi^2 = 2p^{\frac{1}{2}} \xi^2$$

$$\Rightarrow p \geqslant \frac{S^4}{4\xi^4}.$$

定理 2.6.3(Chernoff 定理) 设$(X_n,\ n\geqslant 1)$独立同分布,且 $EX_n<0, P(X_n<0)>0, EX_1^2 e^{tX_1}<\infty\ (\forall\ t>0)$,则

$$\lim_{n\to\infty}\frac{1}{n}\log P\Big(\sum_{i=1}^{n}X_i\geqslant 0\Big)=\log\rho,\quad \rho=\inf_t Ee^{tX_1}.$$

证 令 $Y_n=\sum_{i=1}^{n}X_i$,则 $EY_n<0$,且

$$P(Y_n>0)\geqslant P(X_i>0,\ 1\leqslant i\leqslant n)=P^n(X_1>0)>0.$$

这表明:对$\forall\ n\geqslant 1, Y_n$ 满足定理 2.6.1 中的条件. 令

$$M_n(t)=Ee^{tY_n}=(M(t))^n,$$

则
$$\rho_n=Ee^{\tau_n Y_n}=(Ee^{\tau_n Y_1})^n=(M(\tau_n))^n;$$

$$\rho_n=\inf_t Ee^{tY_n}=\inf_t (M(t))^n=(\inf_t M(t))^n=\rho^n$$

$$\Rightarrow M(\tau_n)=M(\tau)$$

$$\Rightarrow \tau_n=\tau,\quad \rho_n=\rho^n\quad (\forall\ n\geqslant 1),\ \rho=M(\tau).$$

现在类似于由 Y 产生变量 Z,而由 Y_n 产生 Z_n,即 r.v. Z_n 具有概率分布

$$P_{Z_n}(dx)=\frac{1}{\rho_n}e^{\tau_n x}P_{Y_n}(dx)=\frac{1}{\rho^n}e^{\tau x}P_{Y_n}(dx).$$

Z_n 的矩生成函数 $M_{Z_n}(t)$为

$$M_{Z_n}(t)=\frac{1}{\rho_n}M_n(\tau+t)=\left(\frac{M(\tau+t)}{\rho}\right)^n.$$

设 V_i 是由 X_i 产生的 r.v. 则$(V_i,\ 1\leqslant i\leqslant n)$独立同分布,且 Z_n 与 $\sum_{i=1}^{n}V_i$ 具有相同的矩生成函数. 按唯一性,Z_n 与 $\sum_{i=1}^{n}V_i$ 具有相同的概率分布,于是

$$EZ_n=E\Big(\sum_{i=1}^{n}V_i\Big)=0.$$

$$S_n^2=EZ_n^2=E\Big(\sum_{i=1}^{n}V_i\Big)^2=nEV_1^2=nS_1^2=\frac{n}{\rho}M''(\tau)>0.$$

$$\xi_n^4 = EZ_n^4 = E\left(\sum_{i=1}^{n} V_i\right)^4 = \sum_{i_1 i_2 i_3 i_4} E(V_{i_1} V_{i_2} V_{i_3} V_{i_4})$$
$$= nEZ_1^4 + 3n(n-1)(EZ_1^2)^2$$
$$= n\xi_1^4 + 3n(n-1)S_1^4.$$

最后的这个计算公式可以从图 2.6-1 中得出.

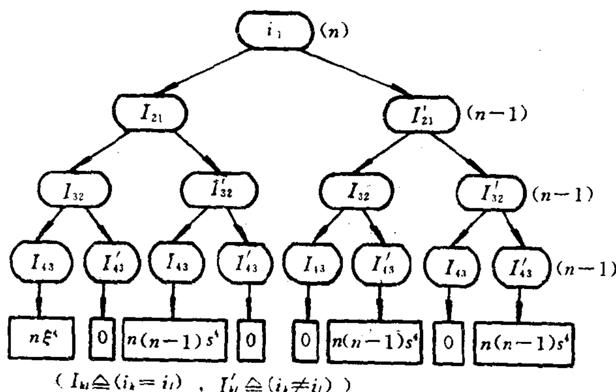


图 2.6-1 ξ_n^4 的计算图

(注意:图 2.6-1 中相应的结果均要用到$(V_i,\ 1 \leqslant i \leqslant n)$的独立性)

上述分析表明:Z_n 满足定理 2.6.2 中的条件. 于是

$$P(Z_n \geqslant 0) \geqslant \frac{S_n^4}{4\xi_n^4} = \frac{n^2 S_1^4}{4(n\xi_1^4 + 3n(n-1)S_1^4)}.$$

按假设 $\xi_1^4 < \infty$. 于是,$\exists\ \alpha > 0$(与 n 无关),使 $\dfrac{S_n^4}{4\xi_n^4} \geqslant \alpha\ (\forall\ n \geqslant 1)$. 按定理 2.6.1,有

$$P(Y_n \geqslant 0) = \rho^n e^{-\theta_n},$$

$$0 < \theta_n \leqslant \frac{\tau S_n}{P(Z_n \geqslant 0)} - \log P(Z_n \geqslant 0) \leqslant \frac{\tau S_n}{\alpha} - \log\alpha$$

$$= \frac{\tau\sqrt{n}S_1}{\alpha} - \log\alpha = O(\sqrt{n})$$

$$\Rightarrow \frac{1}{n}\log P(Y_n \geqslant 0) = \frac{1}{n}(n\log\rho - O(\sqrt{n})) = \log\rho - O\left(\frac{1}{\sqrt{n}}\right)$$

$$\Rightarrow \lim_{n\to\infty}\frac{1}{n}\log P(Y_n \geqslant 0) = \log\rho.$$

Chernoff 定理在假设检验问题中有重要的应用. 设$(X_n,\ n\geqslant 1)$独立同分布,取值于$\{0,1\}$,分布为

$$P(X_1 = 1) = p,\quad P(X_1 = 0) = q = 1 - p.$$

令 $S_n=\sum_{i=1}^{n}X_i$,则

$$P(S_n \geqslant na) = P\left(\sum_{k=1}^{n}(X_k - a) \geqslant 0\right).$$

假设 $0<p<a<1$, $Y_k=X_k-a$,则

$$EY_k = (1-a)p - aq = p - a < 0.$$

$$P(Y_k > 0) = P(X_k > a) = P(X_k = 1) = p > 0.$$

这表明:当 $0<p<a<1$ 时,$(Y_n,\ n\geqslant 1)$满足 Chernoff 定理中的要求.

首先求 $\log\rho$. 计算 Y_k 的矩生成函数:

$$M(t) = Ee^{tY_k} = pe^{t(1-a)} + qe^{-ta}.$$

让 $M'(t)=0$,即得:$e^{\tau}=\frac{aq}{bp}$,其中 $b=1-a$. 于是

$$\rho = M(\tau) = p\left(\frac{aq}{bp}\right)^{1-a} + q\left(\frac{aq}{bp}\right)^{-a} = \frac{q}{b}\left(\frac{bp}{aq}\right)^{a};$$

$$\log\rho = \log\left(\frac{q}{b}\right) + a\left(\log\left(\frac{p}{a}\right) - \log\left(\frac{q}{b}\right)\right)$$

$$= a\log\left(\frac{p}{a}\right) + b\log\left(\frac{q}{b}\right);$$

$$-\log\rho = a\log\left(\frac{a}{p}\right) + b\log\left(\frac{b}{q}\right).$$

按 Chernoff 定理,得

$$\lim_{n\to\infty}\frac{1}{n}\log P\left(\sum_{k=1}^{n}Y_k \geqslant 0\right) = \log\rho$$

$$\Rightarrow P(S_n \geqslant na) \approx \exp(n\log\rho).$$

令 $K^{*}(a,p) = -\log\rho = a\log\left(\frac{a}{p}\right) + b\log\left(\frac{b}{q}\right)$,则

$$P(S_n \geqslant na) \approx \exp(-nK(a,p)) \text{（当 } n \text{ 足够大时）.}$$

现在考虑假设检验问题：假定参数 p 未知，即不知试验序列 $(X_n,\ n\geqslant 1)$ 的分布. 但分布只依赖于一个参数 p. 关于参数 p 有两个竞争性的假设.

假设 H_1：$p=p_1$； 假设 H_2：$p=p_2$.

其中 $0<p_1<p_2<1$.

情形 A 设 $0<p_1<a<p_2<1$，$(X_k,\ 1\leqslant k\leqslant n)$ 为 Bernoulli 序列的 n 次试验结果，让 $S_n=\sum_{i=1}^{n} X_i$.

如果 $S_n\geqslant na$，就接受假设 H_2.

如果 $S_n<na$，就接受假设 H_1.

让 $P(S_n\geqslant na\,|\,H_1)$ 表示接受假设 H_1 的错误概率. $P(S_n<na\,|\,H_2)$ 表示接受假设 H_2 的错误概率. 按假定，

$$P(S_n \geqslant na\,|\,H_1) \sim \exp(-nK(a,p_1)),\ \text{当 } n\to\infty \text{ 时；}$$

$$P(S_n < na\,|\,H_2) \sim \exp(-nK(a,p_2)),\ \text{当 } n\to\infty \text{ 时.}$$

事实上，

$$P(S_n < na\,|\,H_2) = P\Big(\sum_{k=1}^{n}(a-X_k) > 0\,|\,H_2\Big)$$

$$= P_2\Big(\sum_{k=1}^{n}(a-X_k) > 0\Big),$$

其中 P_2 表示对应参数 p_2 的概率测度. 令 $\widetilde{Y}_k=a-X_k$，则

$$E\widetilde{Y}_k = p_2\cdot(a-1)+q_2\cdot a = a-p_2 < 0;$$

$$P_2(\widetilde{Y}_k > 0) = P_2(a-X_k>0) = P_2(X_k=0) = q_2 > 0.$$

故按 Chernoff 定理，有

$$P\Big(\sum_{k=1}^{n}\widetilde{Y}_k \geqslant 0\,|\,H_2\Big) \sim \exp(-nK(a,p_2)) \text{（当 } n\to\infty \text{ 时）.}$$

情形 B 当 $0<p_1<p_2<1$ 时，给定 $a\in(p_1,p_2)$，使得

$$K(a,p_1) = K(a,p_2), \tag{2.6-1}$$

即两个偏差概率相等. 问在此种情形下，会有些什么样的结果？方

程(2.6-1)即为

$$a\log\left(\frac{a}{p_1}\right)+b\log\left(\frac{b}{q_1}\right)=a\log\left(\frac{a}{p_2}\right)+b\log\left(\frac{b}{q_2}\right). \quad (2.6\text{-}2)$$

解方程(2.6-2),即可得解

$$a(p_1,p_2)=\frac{\log\left(\frac{q_1}{q_2}\right)}{\log\left(\frac{p_2}{p_1}\right)+\log\left(\frac{q_1}{q_2}\right)}.$$

$$K(a(p_1,p_2),\ p_1)=a\log\left(\frac{a}{p_1}\right)+(1-a)\log\left(\frac{1-a}{1-p_1}\right).$$

考虑下列估计:

$$\log(1+x)=x-\frac{1}{2}x^2+O(x^3)\quad 当\ x\to 0\ 时,$$

$$K(p+x,p)=(p+x)\log\left(\frac{p+x}{p}\right)$$
$$+(1-p-x)\log\left(\frac{1-p-x}{1-p}\right);$$

$$\log\left(1+\frac{x}{p}\right)=\frac{x}{p}-\frac{1}{2}\frac{x^2}{p^2}+O(x^3);$$

$$\log\left(1-\frac{x}{q}\right)=-\frac{x}{q}-\frac{1}{2}\frac{x^2}{q^2}+O(x^3);$$

$$(p+x)\log\left(1+\frac{x}{p}\right)=x+\frac{1}{2}\frac{x^2}{p}+O(x^3);$$

$$(q-x)\log\left(1-\frac{x}{q}\right)=-x+\frac{1}{2}\frac{x^2}{q}+O(x^3)$$

$$\Rightarrow K(p+x,p)=\frac{x^2}{2pq}+O(x^3).$$

让 $p_1=p$, $p_2=p+t$,则

$$\psi(t)\triangleq a(p,p+t)=p+\frac{1}{2}t+O(t^2).$$

$$\Rightarrow K(a(p,p+t),p)=K\left(p+\frac{1}{2}t+O(t^2),p\right)$$

$$=\frac{t^2}{8pq}+O(t^3)\quad (\text{当 } t\to 0 \text{ 时}).$$

命题2.6.1 为了对很小的 t 值能区分 $p_1=p$ 和 $p_2=p+t$，在选取 a 使两个偏差概率相等的前提下，偏差概率大约为 $\exp\left(-\frac{nt^2}{8pq}\right)$.

当 $p=\frac{1}{2}$ 时，$pq=\frac{1}{4}$（最大值），偏差率为 $\exp\left(-\frac{nt^2}{2}\right)$；

当 $p<\frac{1}{2}$ 时，偏差率为 $\exp\left(-\frac{nt^2}{2pq}\right)<\exp\left(-\frac{nt^2}{2}\right)$.

例如，设 $p_1=0.1$. 这时，有

$$\exp\left(-\frac{nt^2}{8\times 0.1\times 0.9}\right)=\exp\left(-\frac{nt^2}{0.72}\right)<\exp\left(-\frac{nt^2}{2}\right).$$

注意，$\frac{nt^2}{2}=0.36\cdot\frac{nt^2}{0.72}$. 这表明：当 0.5 与 $0.5+t$ 能够区分开时，0.1 与 $0.1+t$ 能区分开只要求有 $36\%n$ 的样本长度.

注 如果偏差概率 $\exp(-nK(a(p_1,p_2),p_1))$ $(0<p_1<p_2<1)$ 很小，则在分布 p_1 下接受假设 H_1；在分布 p_2 下接受假设 H_2.

（三）重对数律

定理 2.6.4 设 $(X_n,\ n\geqslant 1)$ 独立同分布. $EX_k=0$，$EX_k^2=1$，$EX_1^2e^{tX_1}<\infty$ $(\forall\ t>0)$，$S_n=\sum\limits_{k=1}^{n}X_k$. 若常数列 $(a_n,\ n\geqslant 1)$ 满足要求：$a_n\to\infty$，$\frac{a_n}{\sqrt{n}}\xrightarrow[n\to\infty]{}0$，则

$$P\left(S_n\geqslant\sqrt{n}\,a_n\right)=\exp\left(-\frac{a_n^2}{2}(1+\xi_n)\right),\text{其中 } \xi_n\to 0.$$

证 令 $Y_n=S_n-\sqrt{n}\,a_n$，则 $EY_n=-\sqrt{n}\,a_n<0$ 且

$$P(Y_n>0)\geqslant P^n\left(\left\{X_1-\frac{a_n}{\sqrt{n}}>0\right\}\right)>0\quad(\text{当 } n \text{ 足够大时}).$$

于是，当 n 足够大时，定理 2.6.1 中的条件满足. 设 $M_n(t)$，$C_n(t)$，τ_n，ρ_n，Z_n 均与 Y_n 相应的量（见定理 2.6.1），则有

$$P(Y_n \geqslant 0) = \rho_n e^{-\theta_n}, \quad e^{-\theta_n} = E(e^{-\tau_n Z_n} 1_{\{Z_n \geqslant 0\}}).$$

首先考察 τ_n. 设 $m(t)$, $c(t)$分别为 X_1 的矩生成和累积生成函数,即 $m(t)=Ee^{tX_1}$, $c(t)=\log m(t)$. 按独立性,有

$$M_n(t) = Ee^{tY_n} = \left(m(t)\exp\left(-\frac{ta_n}{\sqrt{n}}\right)\right)^n.$$

$$C_n(t) = \log M_n(t) = -t\sqrt{n}\,a_n + nc(t).$$

显然,τ_n 满足方程 $M'_n(t)=0$,也满足 $C'_n(t)=0$,即

$$c'(\tau_n) = \frac{a_n}{\sqrt{n}}.$$

注意,$c(0)=c'(0)=0$, $c''(0)=m''(0)=EX_1^2=1$. 因此,0 为 $c(t)$的最小值点,且 $\tau_n \to 0$. 由$\lim\limits_{n\to\infty}\frac{c'(\tau_n)}{\tau_n}=c''(0)=1$ 知

$$c'(\tau_n) = \tau_n + O(\tau_n^2) \Rightarrow \tau_n = \frac{a_n}{\sqrt{n}} + O\left(\frac{a_n^2}{n}\right).$$

现在考察 ρ_n

$$\begin{aligned}
\log\rho_n &= \log M_n(\tau_n) = -\tau_n\sqrt{n}\,a_n + nc(\tau_n) \\
&= -a_n\sqrt{n}\,\tau_n + n\int_0^{\tau_n} c'(t)\,dt \\
&= -a_n\sqrt{n}\,\tau_n + n\int_0^{\tau_n} (t + O(t^2))\,dt \\
&= -\frac{1}{2}a_n^2 + O(a_n^2).
\end{aligned}$$

最后考察 θ_n. 辅变量 Z_n 具有矩生成函数$\frac{1}{\rho_n}M_n(\tau_n+t)$及累积生成函数

$$\begin{aligned}
D_n(t) &= C_n(\tau_n + t) - \log\rho_n \\
&= -(\tau_n + t)a_n\sqrt{n} + nC(\tau_n + t) - \log\rho_n.
\end{aligned}$$

注意,Z_n 的均值为 0,即$\frac{1}{\rho_n}M'_n(\tau_n)=0$. 从而,有 $D'_n(0)=0$.

$$S_n^2 = EZ_n^2 = D''_n(0) = nC''(\tau_n) = n(C''(0) + O(\tau_n))$$

$$=n(1+O(\tau_n));$$

$$D_n^4(0)=nC^{(4)}(\tau_n)=O(n);$$

$$EZ_n^4=3S_n^4+D_n^{(4)}(0)=3n(1+o(1))+D_n^{(4)}(0)$$

$$\Rightarrow\frac{EZ_n^4}{S_n^4}\xrightarrow[n\to\infty]{}3+\lim_{n\to\infty}\frac{D_n^{(4)}(0)}{S_n^4}.$$

这表明:∃ $\alpha>0$,于 α 与 n 无关,且$\dfrac{S_n^4}{4EZ_n^4}\geqslant\alpha$.

按定理 2.6.2,有

$$P(Z_n\geqslant 0)\geqslant\alpha>0.\ \text{当 } n \text{ 足够大时.}$$

按定理 2.6.1,有

$$\tau_nS_n<\theta_n\leqslant\tau_nS_n\alpha^{-1}-\log\alpha.$$

按 τ_n,S_n 的估计,得

$$\theta_n=O(a_n)=o(a_n^2).$$

综上分析,得

$$P(Y_n\geqslant 0)=\exp\left(-\frac{1}{2}a_n^2(1+o(1))\right).$$

令 $\xi_n=o(1)$,则 $\xi_n\to 0$. 从而,得所要的结论.

定理 2.6.5 设$(X_n,\ n\geqslant 1)$独立同分布,具有均值 0 和方差 1,则当 $\alpha\geqslant\sqrt{2}$ 时,有

$$P\left(\frac{M_n}{\sqrt{n}}\geqslant\alpha\right)\leqslant 2P\left(\frac{S_n}{\sqrt{n}}\geqslant\alpha-\sqrt{2}\right),$$

其中 $M_n=\bigvee_{k=0}^{n}S_k$, $S_0=0$; $S_k=\sum_{i=1}^{k}X_i$.

证 令 $A_j=\{M_{j-1}<\alpha\sqrt{n}\leqslant M_j\}$,则对∀ $0<\beta\leqslant\alpha$,有

$$\left\{\frac{M_n}{\sqrt{n}}\geqslant\alpha\right\}=\bigcup_{j=1}^{n}A_j=\left(\bigcup_{j=1}^{n}A_j\right)\cap\left\{\frac{S_n}{\sqrt{n}}\geqslant\alpha-\beta\right\}$$

$$\cup\left(\bigcup_{j=1}^{n}A_j\right)\cap\left\{\frac{S_n}{\sqrt{n}}<\alpha-\beta\right\}$$

$$\subset\left(\bigcup_{j=1}^{n-1}A_j\right)\cap\left\{\frac{S_n}{\sqrt{n}}<\alpha-\beta\right\}\cup\left\{\frac{S_n}{\sqrt{n}}\geqslant\alpha-\beta\right\}$$

$$\Rightarrow P\left(\left\{\frac{M_n}{\sqrt{n}} \geqslant \alpha\right\}\right) \leqslant P\left(\frac{S_n}{\sqrt{n}} \geqslant \alpha - \beta\right)$$

$$+ \sum_{j=1}^{n-1} P\left(A_j \cap \left\{\frac{S_n}{\sqrt{n}} \leqslant \alpha - \beta\right\}\right).$$

注意,

$$A_j \cap \left\{\frac{S_n}{\sqrt{n}} \leqslant \alpha - \beta\right\}$$

$$= A_j \cap \left\{\frac{S_j}{\sqrt{n}} \geqslant \alpha\right\} \cap \left\{\frac{S_n}{\sqrt{n}} \geqslant \alpha - \beta\right\}$$

$$= A_j \cap \left\{\frac{S_j}{\sqrt{n}} \geqslant \alpha\right\} \cap \left\{\frac{S_n - S_j}{\sqrt{n}} \geqslant \alpha - \frac{S_j}{\sqrt{n}} - \beta\right\}$$

$$\subset A_j \cap \left\{\frac{S_n - S_j}{\sqrt{n}} \leqslant -\beta\right\} \subset A_j \cap \left\{\frac{|S_n - S_j|}{\sqrt{n}} \geqslant \beta\right\},$$

并考虑$(X_n, n \geqslant 1)$的独立性,有

$$P\left(A_j \cap \left\{\frac{S_n}{\sqrt{n}} \leqslant \alpha - \beta\right\}\right) \leqslant P(A_j) \cdot P\left(\frac{|S_n - S_j|}{\sqrt{n}} \geqslant \beta\right)$$

$$\leqslant \frac{1}{n\beta^2} E|S_n - S_j|^2 P(A_j)$$

$$\leqslant \frac{1}{\beta^2} P(A_j)$$

$$\Rightarrow P\left(\frac{M_n}{\sqrt{n}} \geqslant \alpha\right) \leqslant P\left(\frac{S_n}{\sqrt{n}} \geqslant \alpha - \beta\right) + \frac{1}{\beta^2} \sum_{j=1}^{n} P(A_j)$$

$$\Rightarrow P\left(\frac{M_n}{\sqrt{n}} \geqslant \alpha\right) \leqslant \frac{\beta^2}{\beta^2 - 1} P\left(\frac{S_n}{\sqrt{n}} \geqslant \alpha - \beta\right)$$

当 $1 < \beta \leqslant \alpha$ 时.

选取 $\beta = \sqrt{2} \leqslant \alpha$,则由上式即得所要的结论.

注 这个定理可以推广成如下形式:

若 $1 < \beta \leqslant \alpha$,则

$$P\left(\frac{M_n}{\sqrt{n}} \geqslant \alpha\right) \leqslant \frac{\beta^2}{\beta^2 - 1} P\left(\frac{S_n}{\sqrt{n}} \geqslant \alpha - \beta\right).$$

定理 2.6.6(重对数律)　设$(X_n, n\geqslant 1)$独立同分布,具有均值 0 和方差 1. $S_0=0$, $S_n=\sum_{k=1}^{n}X_k$. 则

$$P\left(\overline{\lim_{n\to\infty}}\frac{S_n}{\varphi(n)}=1\right)=1, \text{其中 } \varphi(n)=(2n\log\log n)^{\frac{1}{2}}.$$

此结论等价于(i)和(ii)同时成立:

(i) $P(S_n\geqslant(1+\varepsilon)\varphi(n), \text{i.o})=0\ (\forall\ \varepsilon>0)$;

(ii) $P(S_n\geqslant(1-\varepsilon)\varphi(n), \text{i.o})=1\ (\forall\ \varepsilon>0)$.

证　按定理 2.6.5,对$\forall\ 1<\beta_n\leqslant\theta_n x_n$,有

$$P\left(\frac{M_n}{\sqrt{n}}\geqslant\theta_n x_n\right)\leqslant\frac{\beta_n^2}{\beta_n^2-1}P\left(\frac{S_n}{\sqrt{n}}\geqslant\theta_n x_n-\beta_n\right).$$

若 $x_n\uparrow\infty$, $\frac{x_n}{\sqrt{n}}\to 0$,则按定理 2.6.4,有

$$P\left(\frac{S_n}{\sqrt{n}}\geqslant\theta_n x_n-\beta_n\right)=\exp\left(-\frac{1}{2}(\theta_n x_n-\beta_n)^2(1+\xi_n)\right),$$

其中 $\xi_n\to 0$.

选正数列$(\delta_k, k\geqslant 1)$, 使 $\delta_k\downarrow 0$, $\sum_k\delta_k<\infty$. 然后选正整数列$(n_k, k\geqslant 1)$, 使得

$$\theta_{n_k}x_{n_k}-\beta_{n_k}\geqslant\left(\frac{2}{1+\xi_{n_k}}\log\left(\frac{\beta_{n_k}^2}{\beta_{n_k}^2-1}\delta_k^{-1}\right)\right)^{\frac{1}{2}}.$$

则按 Borel-Cantelli 引理,有

$$P\left(\frac{M_{n_k}}{\sqrt{n_k}}\geqslant\theta_{n_k}x_{n_k}, \text{ i.o.}\right)=0.$$

选 $\beta_n=\beta>1$,令 $\beta^*=\frac{\beta^2}{\beta^2-1}>1$,则

$$\theta_{n_k}x_{n_k}\geqslant(2\log\delta_k^{-1}+2\log\beta^*)^{\frac{1}{2}}+\beta.$$

令 $\nu_k=\frac{\sqrt{n_{k-1}}x_{n_{k-1}}}{\sqrt{n_k}x_{n_k}}$,则当 n 满足 $n_{k-1}<n\leqslant n_k$ 时,有

$$\left\{\frac{S_n}{\sqrt{n}}>(1+\varepsilon)x_n\right\}\subset\left\{\frac{S_n}{\sqrt{n_{k-1}}}>(1+\varepsilon)x_{n_{k-1}}\right\}$$

$$=\left\{S_n>(1+\varepsilon)\nu_k\sqrt{n_k}\,x_{n_k}\right\}\subset\left\{\frac{M_{n_k}}{\sqrt{n_k}}>(1+\varepsilon)\nu_k x_{n_k}\right\}.$$

令 $\theta_{n_k}\leqslant(1+\varepsilon)\nu_k$，则

$$P\left(\frac{S_n}{\sqrt{n}}>(1+\varepsilon)x_n,\ \mathrm{r.o.}\right)=0.$$

为证结论(i)，取

$$x_n=(2\log\log n)^{\frac{1}{2}},\ \log n_k=\left(\frac{1}{q}\right)^k\quad(0<q<1),$$

则 $x_{n_k}=\left(2\log\left(\frac{1}{q}\right)^k\right)^{\frac{1}{2}}$, $n_k=\exp\left(q^{-k}\right)$, $\nu_k=\left(\frac{k-1}{k}\right)^{\frac{1}{2}}$

$\cdot\exp\left(-\frac{1}{2}(1-q)q^{-k}\right)$. 显然，$\nu_k<1$，且 $\nu_k\to1$. 选取 $\varepsilon'>0$，使得

$$1<\theta_{n_k}=(1+\varepsilon)(1-\varepsilon')\leqslant(1+\varepsilon)\nu_k(\text{当 } k \text{ 足够大时}).$$

于是，选定 $0<q<1$. 让 $\delta_k=q^k$，则当 $\beta>1$ 时，对足够大的 k，有

$$\theta_{n_k}x_{n_k}>(2\log\delta_k^{-1}+2\log\beta^*)^{\frac{1}{2}}+\beta$$

$$\Rightarrow P\left(\frac{S_n}{\sqrt{n}}>(1+\varepsilon)(2\log\log n)^{\frac{1}{2}},\ \mathrm{i.o.}\right)=0\Rightarrow\text{(i) 成立.}$$

下证(ii). 为此，考察 $\varphi(n)$ 应满足哪些要求.

(Ⅰ) 要求$\exists$ 子列$(n_k,\ k\geqslant1)$，使得

$$\frac{\varphi(n_k)}{\sqrt{n_k-n_{k-1}}}\to\infty,\quad\frac{\varphi(n_k)}{n_k-n_{k-1}}\to\infty,$$

$$\exp\left(-\frac{1}{2}\frac{\alpha^2\varphi^2(n_k)}{n_k-n_{k-1}}\right)\geqslant\delta_k,$$

其中 $\delta_k\downarrow0$，$\sum\limits_k\delta_k=\infty$，$0<\alpha<1$. 这时，按定理 2.6.4，有

$$P(S_{n_k}-S_{n_{k-1}}\geqslant \alpha\varphi(n_k))=P(S_{n_k-n_{k-1}}\geqslant \alpha\varphi(n_k))$$
$$=\exp\left(-\frac{1}{2}\frac{\alpha^2\varphi^2(n_k)}{n_k-n_{k-1}}(1+\xi_k)\right),$$
其中 $\xi_k\to 0$.

按 Borel-Cantelli 引理,得

$$P(S_{n_k}-S_{n_{k-1}}\geqslant \alpha\varphi(n_k)\text{; i. o. })=1.$$

按对称性及结论(i),对$\forall\ \theta>1$,有

$$P(-S_{n_{k-1}}\leqslant \theta\varphi(n_{k-1}),\ \text{i. o.})=1$$
$$\Rightarrow P(S_{n_k}\geqslant \tilde{\alpha}\,\varphi(n_k),\ \text{i. o.})$$
$$=P(S_{n_k}-S_{n_{k-1}}\geqslant \tilde{\alpha}\,\varphi(n_k)-S_{n_{k-1}}$$
$$-S_{n_{k-1}}\leqslant \theta\varphi(n_{k-1}),\ \text{i. o.})$$
$$\geqslant P(S_{n_k}-S_{n_{k-1}}\geqslant \tilde{\alpha}\,\varphi(n_k)+\theta\varphi(n_{k-1}),\ \text{i. o.}),$$

其中 $\tilde{\alpha}=1-\varepsilon$.

(Ⅱ)要求 $u>1$,使得$\frac{\varphi(n_{k-1})}{\varphi(n_k)}\leqslant\frac{1}{u}$,令 $\alpha=\tilde{\alpha}+\frac{\theta}{u}$. 并选 $\theta>1$,予 $0<\alpha<1$, 则按(Ⅰ),有

$$P(S_{n_k}-S_{n_{k-1}})\geqslant\tilde{\alpha}\,\varphi(n_k)+\theta\varphi(n_{k-1}),\ \text{i. o.})$$
$$\geqslant P\left(S_{n_k}-S_{n_{k-1}}\geqslant\left(\tilde{\alpha}+\frac{\theta}{u}\right)\varphi(n_k),\ \text{i. o.}\right)=1.$$
$$\Rightarrow P(S_{n_k}\geqslant(1-\varepsilon)\varphi(n_k),\ \text{i. o.})=1.$$

(Ⅲ)验证 $\varphi(n)=(2n\log\log n)^{\frac{1}{2}}$满足要求(Ⅰ)和(Ⅱ).

注意,$\frac{\varphi^2(n_k)}{n_k-n_{k-1}}\leqslant\frac{2}{\alpha^2}\log\left(\frac{1}{\delta_k}\right)$. 取 $n_k=q^k(1<q)$,$\delta_k=\frac{1}{k}$,及 $1-\frac{1}{q}<\alpha^2$. 选 $u>1$,使得

$$\frac{\varphi(n_{k-1})}{\varphi(n_k)}=\left(\frac{1}{q}\ \frac{\log(k-1)+\log\log q}{\log k+\log\log q}\right)^{\frac{1}{2}}\leqslant\frac{1}{u}.$$

为此,u 满足 $1<u^2\leqslant q$,即保证上式成立.

归纳起来得如下一组关系：

$$1<\frac{1}{q}+\alpha^2,\quad \alpha=1-\varepsilon+\frac{\theta}{u},\quad 1<u^2\leqslant q,$$

$$1<\theta,\quad 0<\alpha<1.$$

只要不等式 $1-\varepsilon+\frac{1}{\sqrt{q}}\geqslant\left(1-\frac{1}{q}\right)^2$ 有解 $q>1$，则上述关系组就有解存在. 然而，当 $\varepsilon>0$ 足够小时，此不等式一定有这种解存在.

第三章　条件期望

条件期望理论是现代概率论中的基本内容之一. 它在过程论中起着特别重要的作用. Markov 过程,鞅过程等都是在此理论基础上提出并展开的. 本章着重介绍条件期望的基本概念,基本性质及某些应用. 条件概率将作为条件期望的特例仅作简要介绍.

§3.1　Radon-Nikodym 定理

R.-N. 定理是测度论中极其重要的结论之一. 它是条件期望理论发展的基本根据. 本节将比较详细地介绍这方面的知识.

(一) Hahn 分解和 Jordan 分解

定义 3.1.1　设$(\Omega,\mathscr{F})$为可测空间,ν 为 $\mathscr{F}$ 上的实值集函数,且具有 σ-可加性,则称 ν 为$(\Omega,\mathscr{F})$上的有限广义测度.

按此定义,有限广义测度 ν 具有下列性质:

(i) $\nu(\varnothing)=0$, ν 具有可加性;

(ii) $\nu(A\backslash B)=\nu(A)-\nu(B)$,其中 $A,B\in\mathscr{F}$, $B\subset A$.

定义 3.1.2　设 ν 为$(\Omega,\mathscr{F})$上的有限广义测度,$E\in\mathscr{F}$. 若对 $\forall\ A\in\mathscr{F}$,且 $A\subset E$,有 $\nu(A)\geqslant 0(\leqslant 0)$,则称 E 为 ν 的**正集(负集)**.

按此定义,正、负集具有下列特点:

(i) 若 $E\in\mathscr{F}$ 同时为正、负集,则 $\nu(A\cap E)=0$ $(\forall\ A\in\mathscr{F})$;

(ii) 若 $E\in\mathscr{F}$ 为正集,则对$\forall\ A\in\mathscr{F}$,$A\cap E$ 亦为正集,且 $\nu(A\cap E)\leqslant\nu(E)$,即 ν 为 $\mathscr{F}\cap E$ 上的一个有限测度.

引理 3.1.1　有限个正(负)集之交,可列个正(负)集之并,两个正(负)集之差均为正(负)集.

证　设 E_1、E_2 为两个正集,则 $E_1\cap E_2\subset E_1$, $E_1\backslash E_2\subset E_1$, $E_1\cup$

$E_2=E_1\backslash E_2\cup E_2\backslash E_1\cup E_1\cap E_2$. 从而，由 ν 的 σ-可加性知：结论对有限个正集的情形是正确的. 若 $(E_k,\ k\geqslant 1)\subset\mathscr{F}$，且每个 E_k 为正集，则

$$\bigcup_k E_k=\bigcup_k \widetilde{E}_k,$$

其中 $\widetilde{E}_1=E_1$, $\widetilde{E}_k=E_k\backslash\bigcup_{i=1}^{k-1}E_i=\bigcap_{i=1}^{k-1}(E_k\backslash E_i)$ $(k\geqslant 2)$. 且每个 $\widetilde{E}_k$ 为正集. 按 ν 的 σ-可加性，$\bigcup_k \widetilde{E}_k$ 为正集.

定义 3.1.3 若 E 为 ν 的负集，且 $\nu(E)<0$，则称 E 为 ν 的真负集.

按此定义，真负集的可测子集不必为真负集.

引理 3.1.2 若 $A\in\mathscr{F}$ 为不含 ν 的真负子集，则 A 为正集.

证 采用反证法. 假若 A 不为正集(但也不必为负集)，则 $\exists\ E_0\subset A$, $E_0\in\mathscr{F}$, $\ni\ \nu(E_0)<0$. 按条件，E_0 不为真负集. 因此，E_0 不为负集. 令

$$K_0=\sup\{\nu(E):\ E\subset E_0,\ E\in\mathscr{F}\},$$

则 $K_0>0$. 于是，$\exists\ E_1\subset E_0$, $E_1\in\mathscr{F}$, $\ni\ \nu(E_1)>\frac{1}{2}K_0$. 这时，$E_0\backslash E_1\subset A$，且 $\nu(E_0\backslash E_1)=\nu(E_0)-\nu(E_1)<0$. 按条件，$E_0\backslash E_1$ 也不是负集. 以 $E_0\backslash E_1$ 代替 E_0，重复上述论证过程，则可得 $K_1>0$, $E_2\subset E_0\backslash E_1$, $\nu(E_2)>\frac{1}{2}K_1$, $\nu(E_0\backslash E_1\backslash E_2)<0$, $E_0\backslash E_1\backslash E_2$ 不为负集. 继续这一构作过程，便可得一数例 $(K_n,\ n\geqslant 0)$，集列 $(E_n,\ n\geqslant 0)$，具有下列性质：

(1) $\bigcup_n E_n\subset E_0$, $E_n=E_0\backslash\bigcup_{k=1}^{n-1}E_k$ $(n\geqslant 2)$, $E_n\cap E_m=\varnothing$ $(n\neq m)$;

(2) $K_n=\sup\{\nu(E):\ E\subset E_0\backslash\bigcup_{k=1}^{n}E_k\}>0$;

(3) $\nu(E_n)>\frac{1}{2}K_{n-1}(n\geqslant 1)$, $\nu(E_0)<0$, $\nu(E_0\backslash\bigcup_{k=1}^{n}E_k)<0$;

(4) $K_n\to 0$ $\left(\because\ 0<\frac{1}{2}\sum_{n=1}^{\infty}K_{n-1}<\sum_{n=1}^{\infty}\nu(E_n)=\nu(\bigcup_{n=1}^{\infty}E_k)<\infty\right)$;

(5) $\nu(F_0)=\nu(E_0)-\nu(\bigcup_{k=1}^{\infty}E_k)<0$，其中 $F_0=E_0\setminus\bigcup_{k=1}^{\infty}E_k\subset A$；

(6) 对$\forall$ $F_0\supset A'\in\mathscr{F}$，有

$$\nu(A')\leqslant\sup\{\nu(E):E\subset E_0\setminus\bigcup_{k=1}^{n}E_k\}=K_n\to 0$$

$$\Rightarrow F_0 \text{ 为 } A \text{ 中的负集.}$$

性质(5)和(6)表明：F_0 为 A 中的真负集. 这与假设矛盾. 故引理成立.

前面提到过，ν 为它的正集之可测子集类上的测度，那么是否在 Ω 中存在一个最大的可测正子集呢？下面的分解定理就是回答这个问题的.

定理 3.1.1(Hahn 分解)　设 ν 为$(\Omega,\mathscr{F})$上的有限广义测度，则存在 ν 的正集 A 和负集 B，使得

$$A\cup B=\Omega,\quad A\cap B=\varnothing.$$

证　令 $\alpha=\sup\{\nu(E):E$ 为正集$\}$，则可选可测正集列$(A_n,n\geqslant1)$，于 $\lim\limits_{n\to\infty}\nu(A_n)=\alpha$，令 $A=\bigcup\limits_{n=1}^{\infty}A_n$，则 A 具有下列性质：

(1) A 为正集(按引理 3.1.1)；

(2) $\nu(A)=\alpha$. 事实上，按(1)，有 $\nu(A)\leqslant\alpha$. 按 ν 在正集类上的单调性，有

$$\nu(A)\geqslant\nu(\bigcup_{k=1}^{n}A_k)\to\alpha\Rightarrow\nu(A)=\alpha;$$

(3) 令 $B=\Omega\setminus A=A^c$，则 B 为 ν 的负集. 事实上，按引理 3.1.2(将负换成正)，仅需证明：B 中不含可测真正子集. 假若不然，B 含有一个真正集 $E_0\in\mathscr{F}$，即 E_0 为正集，且 $\nu(E_0)>0$. 那么，由 $E_0\cap A=\varnothing$ 而有

$$\nu(A\cup E_0)=\nu(A)+\nu(E_0)>\alpha.$$

这不可能，因 $A\cup E_0$ 为正集.

归纳上述分析，得分解

$$\Omega=A\cup B,\ A\cap B=\varnothing,\ A,B \text{ 分别为正、负集.}$$

注　Hahn 分解一般不唯一. 原因在于：有可能存在非空可测

子集 D,使之同时为 ν 的正、负集. 于是,$\nu(D)=0$. 这时,将 D 加到 A 或 B 中均不改变它们的正、负性. 然而,这个事实表明:Hahn 分解可唯一到一个 ν-零集.

定义 3.1.4 设 ν 为 $(\Omega, \mathscr{F})$ 上的有限广义测度,A、$B\in\mathscr{F}$ 为关于 ν 的 Hahn 分解. 定义测度 $\nu^{\pm}$:

$$\nu^+(E)=\nu(E\cap A),\ \nu^-(E)=-\nu(E\cap B)\quad(\forall\ E\in\mathscr{F}).$$

则称 ν^+, ν^- 分别为 ν 的上,下变差,称 $|\nu|=\nu^++\nu^-$ 为 ν 的**全变差**.

引理 3.1.3(Jordan 分解). 设 ν 为 $(\Omega,\mathscr{F})$ 上的有限广义测度,则 $\nu=\nu^+-\nu^-$,即

$$\nu(E)=\nu^+(E)-\nu^-(E)\quad(\forall\ E\in\mathscr{F}).$$

证 设 A,B 为 ν 的 Hahn 分解,则

$$\begin{aligned}\nu(E)&=\nu(E\cap A)+\nu(E\cap B)\\&=\nu^+(E)-\nu^-(E)\quad(\forall\ E\in\mathscr{F}).\end{aligned}$$

(二) R.-N. 定理

定义 3.1.5 设 μ,ν 分别为 $(\Omega,\mathscr{F})$ 上的测度和有限广义测度. 若 $E\in\mathscr{F}$, $\mu(E)=0$, 必有 $\nu(E)=0$,则称 ν 关于 μ **绝对连续**. 记为 $\nu\ll\mu$. 若$\exists\ N\in\mathscr{F}$, $\ni$ $\mu(N)=0$, $\nu(E\cap N^c)=0$ $(\forall\ E\in\mathscr{F})$, 则称 ν 关于 μ **是奇异的**. 记为 $\nu\perp\mu$.

引理 3.1.4 让 μ,ν 按定义 3.1.5 中所设,则

$$\nu\ll(\perp)\mu\iff\nu^{\pm}\ll(\perp)\mu\iff|\nu|\ll(\perp)\mu.$$

证 设 A,B 为 Ω 的 Hahn 分解(关于 ν). 注意,

$$\mu(E)=0\iff\mu(A\cap E)=0,\ \mu(B\cap E)=0.$$

若 $\nu\ll\mu$, 则当 $\mu(E)=0$ 时,有 $\nu^+(E)=\nu(A\cap E)=0$,

$$\begin{aligned}&\nu^-(E)=-\nu(B\cap E)=0\\&\Rightarrow\nu^{\pm}\ll\mu,\ |\nu|\ll\mu.\end{aligned}$$

反之亦然.

引理 3.1.5 设 $(\Omega,\mathscr{F},\mu)$ 为测度空间,ν_1, ν_2 为 $(\Omega,\mathscr{F})$ 上的有限广义测度. 若 $\nu_i\ll(\perp)\mu$ $(i=1,2)$,则

$$\nu_1 \pm \nu_2 \ll (\perp)\mu.$$

引理 3.1.6 $\nu \ll \mu$，且 $\nu \perp \mu \Longleftrightarrow \nu(E)=0$ $(\forall\ E \in \mathscr{F})$，其中 μ 为测度，ν 为有限广义测度.

引理 3.1.7 设 $(\Omega, \mathscr{F}, \mu)$ 为 σ-有限测度空间，$(\Omega, \mathscr{F}, \nu)$ 为有限测度空间，则存在非负 $\mathscr{F}$-可测函数 f 及 $(\Omega, \mathscr{F})$ 上的有限测度 ν_c 和 ν_s，$\dashv$

$$\nu = \nu_c + \nu_s,$$

其中 $\nu_s \perp \mu$，$\nu_c \ll \mu$，且 $\nu_c(E) = \int_E f \mathrm{d}\mu$ $(\forall\ E \in \mathscr{F})$.

证 首先讨论 μ 为有限测度的情形. 令

$$F = \left\{f:\ f \text{ 非负 } \mathscr{F}\text{- 可测，且} \int_E f \mathrm{d}\mu \leqslant \nu(E)\ (\forall\ E \in \mathscr{F})\right\}.$$

显然，$f=0 \in F$. 因此，F 非空. 设 $(g_i,\ 1 \leqslant i \leqslant n) \subset F$，则

$$g = \max_{1 \leqslant i \leqslant n} g_i \in F.$$

事实上，令 $\tau(\omega) = \inf\{k:\ g(\omega) = g_k(\omega)\}$，则

$$\Omega = \bigcup_{k=1}^{n} (\tau = k), \quad (\tau = k) \cdot (\tau = l) = \varnothing\ (k \neq l)$$

$$\Rightarrow \int_E g \mathrm{d}\mu = \sum_{k=1}^{n} \int_{E \cap (\tau = k)} g \mathrm{d}\mu = \sum_{k=1}^{n} \int_{E \cap (\tau = k)} g_k \mathrm{d}\mu$$

$$\leqslant \sum_{k=1}^{n} \nu(E \cap (\tau = k)) = \nu(E) \Rightarrow g \in F.$$

令 $\alpha = \sup\left\{\int_\Omega f \mathrm{d}\mu:\ f \in F\right\}$，则 $\alpha \leqslant \nu(\Omega)$，且可选 $g_i \in F$，$\dashv$ $\alpha = \lim\limits_{n\to\infty} \int_\Omega g_n \mathrm{d}\mu$. 不妨设 $g_n \uparrow$. 令

$$f = \lim_{n\to\infty} g_n.$$

则按单调收敛定理，有

$$\int_\Omega f \mathrm{d}\mu = \lim_{n\to\infty} \int_\Omega g_n \mathrm{d}\mu = \alpha \leqslant \nu(\Omega).$$

定义测度 ν_c：$\nu_c(E) = \int_E f \mathrm{d}\mu$ $(\forall\ E \in \mathscr{F})$，则 ν_c 为有限测度，且 $\nu_c \ll \mu$. 定义测度 ν_s：$\nu_s = \nu - \nu_c$，则为了得到引理中的结论仅需证明：

$\nu_s \perp \mu$. 为此,令

$$\lambda_n = \nu_s - \frac{1}{n}\mu,$$

则对$\forall\ n \geqslant 1$, λ_n 为有限广义测度. 按 Hahn 分解定理,存在 A_n, B_n 分别为 λ_n 的正、负集,使得

$$\Omega = A_n \cup B_n,\ A_n \cap B_n = \varnothing.$$

按 λ_n 的 Jordan 分解,得

$$\nu_s(E \cap A_n) \geqslant \frac{1}{n}\mu(E \cap A_n),$$

$$\nu_s(E \cap B_n) \leqslant \frac{1}{n}\mu(E \cap B_n) \quad (\forall\ E \in \mathscr{F})$$

$$\begin{aligned}\Rightarrow \nu(E) &= \nu_c(E) + \nu_s(E) = \nu_c(E) + \nu_s(E \cap A_n) + \nu_s(E \cap B_n) \\ &\geqslant \nu_c(E) + \frac{1}{n}\mu(E \cap A_n) + \nu_s(E \cap B_n) \\ &\geqslant \nu_c(E) + \frac{1}{n}\mu(E \cap A_n) \\ &\doteq \int_E h_n \mathrm{d}\mu \quad (\forall\ E \in \mathscr{F}),\end{aligned}$$

其中 $h_n = f + \frac{1}{n}1_{A_n}$, 且 $h_n \in F$. 从而, $\int_\Omega h_n \mathrm{d}\mu \leqslant \alpha$. 注意, $\int_\Omega f \mathrm{d}\mu = \alpha$.

故 $\mu(A_n) = 0\ (\forall\ n \geqslant 1)$. 令 $N = \bigcup_{n=1}^{\infty} A_n$, 则 $\mu(N) = 0$,

$$\nu_s(N^c) = \nu_s(\bigcap_n A_n^c) = \nu_s(\bigcap_n B_n) \leqslant \nu_s(B_n) \leqslant \frac{1}{n}\mu(B_n) \xrightarrow[n\to\infty]{} 0$$

$\Rightarrow \nu_s \perp \mu$.

现在讨论 μ 为 σ-有限的情形. μ 的 σ 有限性表明: $\exists\ E_n \in \mathscr{F}\ (n \geqslant 1)$,使得

$$E_i \cap E_j = \varnothing\ (i \neq j),\ \Omega = \bigcup_{n=1}^{\infty} E_n,\ \mu(E_n) < \infty\ (\forall\ n \geqslant 1).$$

令 $\qquad \mathscr{F}_n = \mathscr{F} \cap E_n,\ \mu_n = \mu|_{\mathscr{F}_n}, \nu_n = \nu|_{\mathscr{F}_n},$

则对$\forall\ n \geqslant 1$, $(E_n, \mathscr{F}_n, \mu_n)$为有限测度空间. 按前一步之证,$\exists$ 非负 $\mathscr{F}_n$-可测函数 f_n,使得

$$\nu_n = \nu_c^n + \nu_s^n,$$

其中 $\nu_s^n \perp \mu_n$, $\nu_c^n(E) = \int_E f_n \mathrm{d}\mu_n (\forall\, E \in \mathscr{F}_n)$.

令
$$f(\omega) = \sum_{n=1}^{\infty} f_n(\omega) 1_{E_n}(\omega),$$

$$\nu_s(E) = \sum_{n=1}^{\infty} \nu_s^n(E \cap E_n),\ \nu_c(E) = \sum_{n=1}^{\infty} \nu_c^n(E \cap E_n) \quad (\forall\, E \in \mathscr{F}),$$

则
$$\nu = \nu_c + \nu_s,$$

其中
$$\nu_s \perp \mu,\ \nu_c(E) = \int_E f \mathrm{d}\mu \quad (\forall\, E \in \mathscr{F}).$$

事实上，显然，f 是非负 $\mathscr{F}$-可测的. 由 $\nu_s^n \perp \mu_n$ 知：$\exists\, N_n \in \mathscr{F}_n$，$\ni$ $\mu_n(N_n)=0$，$\nu_s^n(E_n \backslash N_n)=0$. 令 $N=\bigcup_{n=1}^{\infty} N_n$，则

$$\mu(N) = \sum_{n=1}^{\infty} \mu(N_n) = \sum_{n=1}^{\infty} \mu_n(N_n) = 0;$$

$$\begin{aligned}
\nu_s(N^c) &= \nu_s\Big(\bigcup_{n=1}^{\infty} E_n \backslash \bigcup_{n=1}^{\infty} N_n\Big) = \nu_s\Big(\bigcup_{n=1}^{\infty}\Big(E_n \backslash \bigcup_{n=1}^{\infty} N_n\Big)\Big) \\
&= \nu_s\Big(\bigcup_{n=1}^{\infty}(E_n \backslash N_n)\Big) = \sum_{n=1}^{\infty} \nu_s(E_n \backslash N_n) \\
&= \sum_{n=1}^{\infty} \nu_s^n(E_n \backslash N_n) = 0 \\
&\Rightarrow \nu_s \perp \mu.
\end{aligned}$$

此外，对$\forall\, E \in \mathscr{F}$，有

$$\begin{aligned}
\nu_c(E) &= \sum_{n=1}^{\infty} \nu_c^n(E \cap E_n) = \sum_{n=1}^{\infty} \int_{E \cap E_n} f_n \mathrm{d}\mu_n \\
&= \sum_{n=1}^{\infty} \int_{E \cap E_n} f \mathrm{d}\mu = \int_E f \mathrm{d}\mu;
\end{aligned}$$

$$\nu(E) = \sum_{n=1}^{\infty} \nu(E \cap E_n) = \sum_{n=1}^{\infty} \nu_n(E \cap E_n)$$

$$= \sum_{n=1}^{\infty} (\nu_c^n + \nu_s^n)(E \cap E_n) = \nu_c(E) + \nu_s(E).$$

定理 3.1.2(Lebesque 分解)　设$(\Omega, \mathscr{F}, \mu)$为$\sigma$-有限测度空间. $(\Omega, \mathscr{F}, \nu)$为有限测度空间,则∃ 非负$\mathscr{F}$-可测函数$f$及$(\Omega, \mathscr{F})$上的有限测度$\nu_c$和$\nu_s$,使得

$$\nu = \nu_c + \nu_s,$$

其中$\nu_s \perp \mu$, $\nu_c \ll \mu$,且$\nu_c(E) = \int_E f \mathrm{d}\mu \ (\forall E \in \mathscr{F})$.

此外,ν按此格式的分解是唯一的.

证　按引理 3.1.7,分解存在. 按引理 3.1.6,分解是唯一的.

定理 3.1.3(R.-N. 定理)　设$(\Omega, \mathscr{F}, \mu)$为$\sigma$-有限测度空间,$(\Omega, \mathscr{F}, \nu)$为有限广义测度空间,且$\nu \ll \mu$,则∃ $(\mathrm{mod}\mu)$唯一的$\mathscr{F}$-可测函数f,使得

$$\nu(E) = \int_E f \mathrm{d}\mu \ (\forall E \in \mathscr{F}).$$

证　按 Jordan 分解,$\nu = \nu^+ - \nu^-$,其中$\nu^\pm$为有限测度. 按 Lebesque 分解,存在非负$\mathscr{F}$-可测函数$f^\pm$,使得

$$\nu_c^\pm(E) = \int_E f^\pm \mathrm{d}\mu \ (\forall E \in \mathscr{F}).$$

按假设及唯一性,有$\nu_s^\pm = 0$. 从而,$\nu^\pm = \nu_c^\pm$,即

$$\nu^\pm(E) = \int_E f^\pm \mathrm{d}\mu \ (\forall E \in \mathscr{F}).$$

令$f = f^+ - f^-$,则$\nu(E) = \int_E f \mathrm{d}\mu \ (\forall E \in \mathscr{F})$.

最后,唯一性之证等价于证明:若g是$\mathscr{F}$-可测的,且$\int_E g \mathrm{d}\mu = 0 \ (\forall E \in \mathscr{F})$,则$g = 0 \ (\mathrm{mod}\mu)$. 为证此,令$A = \{\omega: g(\omega) \neq 0\}$, $A_1 = \{g > 0\}$, $A_2 = \{g < 0\}$, $A = A_1 \cup A_2$. 假若$\mu(A) > 0$,则$\mu(A_1) \vee \mu(A_2) > 0$. 假定$\mu(A_1) > 0$,则

$$0 = \int_{A_1} g \mathrm{d}\mu > 0.$$

这就出现矛盾. 故$g = 0 (\mathrm{mod}\mu)$.

注 假如在定理 3.1.3 中的 ν 改为 σ-有限测度，则易证此定理仍然正确. 以后，我们将采用此注.

（三）R.-N. 导数

定义 3.1.6 设 $(\Omega,\mathscr{F},\mu)$ 和 $(\Omega,\mathscr{F},\nu)$ 分别为测度和广义测度空间. 若∃ $\mathscr{F}$-可测函数 g，∋ $\nu(E)=\int_E g\mathrm{d}\mu\ (\forall\ E\in\mathscr{F})$，则称 g 为 ν 关于 μ 的 R.-N. 导数，记为 $\frac{\mathrm{d}\nu}{\mathrm{d}\mu}(\omega)=g(\omega)\ (\forall\ \omega\in\Omega)$.

按定理 3.1.3，当 μ 是 σ-有限测度，而 ν 是 σ-有限测度或有限广义测度时，ν 关于 μ 的 R.-N. 导数 $\frac{\mathrm{d}\nu}{\mathrm{d}\mu}$ 在 $\nu\ll\mu$ 的条件下唯一存在 $(\mathrm{mod}\,\mu)$. 下面的结论类似于积分变量替换.

定理 3.1.4 设 $(\Omega,\mathscr{F},\mu)$，$(\Omega,\mathscr{F},\nu)$ 均为 σ-有限测度空间，且 $\nu\ll\mu$. 若 f 是 $(\Omega,\mathscr{F})\to(\mathrm{R},\mathscr{B})$ 的可测函数，且 $\int_\Omega f\mathrm{d}\nu$ 存在，则 $\int_A f\mathrm{d}\nu=\int_A f\frac{\mathrm{d}\nu}{\mathrm{d}\mu}\mathrm{d}\mu\ (\forall\ A\in\mathscr{F})$.

证 $\int_\Omega f\mathrm{d}\nu$ 存在表明：$\int_\Omega f^\pm\,\mathrm{d}\nu$ 中至少有一个有限. 因此，不妨假定，$f\geqslant 0(\mathrm{mod}\,\nu)$. 按 μ,ν 之 σ-有限性，可选出集列 $(E_n,n\geqslant 1)\subset\mathscr{F}$，使得

$$E_n\cap E_m=\varnothing\,(n\neq m),\ \Omega=\bigcup_n E_n,$$
$$\mu(E_n)\vee\nu(E_n)<\infty\ (\forall\ n\geqslant 1).$$

如果结论在 $\mathscr{F}\cap E_n$ 上成立，则利用单调收敛定理便可知：结论在 $\mathscr{F}$ 上成立. 因此，不妨设 μ,ν 均为有限测度. 令

$$H=\left\{x:\ x\ \text{非负}\ \mathscr{F}\text{-可测},\int_A x\frac{\mathrm{d}\nu}{\mathrm{d}\mu}\mathrm{d}\mu=\int_A x\mathrm{d}\nu\ (\forall\ A\in\mathscr{F})\right\}.$$

则结论成立等价于证明：H 包含所有非负 $\mathscr{F}$-可测函数. 为此，首先考虑下形函数：$x(\omega)=1_B(\omega)\ (B\in\mathscr{F})$. 显然，$x$ 是非负 $\mathscr{F}$-可测的，且

$$\int_A x\mathrm{d}\nu=\int_{AB}\mathrm{d}\nu=\nu(AB)=\int_{AB}\frac{\mathrm{d}\nu}{\mathrm{d}\mu}\mathrm{d}\mu=\int_A x\frac{\mathrm{d}\nu}{\mathrm{d}\mu}\mathrm{d}\mu\ (\forall\ A\in\mathscr{F}).$$

(这里引用了定理 3.1.3). 因此,$x\in H$. 由此立即可得任意简单函数 $x=\sum_{i=1}^{n}a_i 1_{B_i}\in H$.

现设 x 为任意非负 $\mathscr{F}$-可测函数. 对此,存在简单非负 $\mathscr{F}$-可测函数列 $(x_n,\ n\geqslant 1)$, ∋ $x_n\uparrow x(\mathrm{mod}\nu)$. 按单调收敛定理,有

$$\int_A x\mathrm{d}\nu=\lim_{n\to\infty}\int_A x_n\mathrm{d}\nu=\lim_{n\to\infty}\int_A x_n\frac{\mathrm{d}\nu}{\mathrm{d}\mu}\mathrm{d}\mu=\int_A x\frac{\mathrm{d}\nu}{\mathrm{d}\mu}\mathrm{d}\mu\ (\forall\ A\in\mathscr{F}).$$

$\Rightarrow x\in H$.

推论 3.1.1 设 ξ 为 $(\Omega,\mathscr{F},P)$ 上的 r.v.,其中 P 为概率测度,且 P_ξ 是由 ξ 诱导的 $(\mathrm{R},\mathscr{B})$ 上的概率测度. λ 为 $(\mathrm{R},\mathscr{B})$ 上的 Lebesque 测度. 此外,$P_\xi\ll\lambda$. 若 g 为 R 上的 Borel 函数,且 $Eg(\xi)$ 存在,则

$$P(\xi\in B)=\int_B f(t)\lambda(\mathrm{d}t),\ \text{其中}\ f=\frac{\mathrm{d}P_\xi}{\mathrm{d}\lambda},$$

$$\int_{\xi^{-1}(B)}g(\xi)\mathrm{d}P=\int_B g(t)f(t)\lambda(\mathrm{d}t)\ \ (\forall\ B\in\mathscr{B}).$$

证 按定理 3.1.3,$f=\frac{\mathrm{d}P_\xi}{\mathrm{d}\lambda}(\mathrm{mod}\lambda)$唯一存在且有限. 于是,$P_\xi(B)=\int_B f\mathrm{d}\lambda\ (\forall\ B\in\mathscr{B})$. 由此即得第一公式.

按积分变换公式,有

$$\int_{\xi^{-1}(B)}g(\xi)\mathrm{d}P=E(1_B(\xi)g(\xi))=\int_{\mathrm{R}}1_B(x)g(x)P_\xi(\mathrm{d}x)$$

$$=\int_{\mathrm{R}}1_B(x)g(x)f(x)\lambda(\mathrm{d}x)\quad(\forall\ B\in\mathscr{B}).$$

如何确定 R.-N. 导数不论在理论上还是在实际上均是很重要的问题. 下面就一类极具实用性的测度给出 R.-N. 导数的求法. 设 μ,ν 分别为 $(R,\mathscr{B})$ 上的概率测度和 σ-有限测度,且 $\nu\ll\mu$. 按定理 3.1.3,$\frac{\mathrm{d}\nu}{\mathrm{d}\mu}(\mathrm{mod}\mu)$唯一存在且有限. 在第一章中,已经知道 R 上的 σ-代数 $\mathscr{B}$ 可由左开右闭的集类产生. 按测度的唯一性定理,立即可得:

$$f(x)=\frac{\mathrm{d}\nu}{\mathrm{d}\mu}=\lim_{h\downarrow 0}\frac{\nu(x-h,x+h]}{\mu(x-h,x+h]}(\mathrm{mod}\mu).$$

实际上，$f(x)$是积分平均值的极限，即

$$\lim_{h\downarrow 0}\frac{1}{\mu(x-h,x+h]}\int_{(x-h,x+h]}f(t)\mu(\mathrm{d}t)$$
$$=f(x)\ (\mathrm{mod}\mu)(x\in\mathrm{R}).$$

§3.2 条件期望定义·存在与唯一性

设$(\Omega,\mathscr{F},P)$为某概率空间. $\mathscr{G}$ 为 $\mathscr{F}$ 中的子σ-代数. P_*为P到$\mathscr{G}$上的限制，即$P_*(A)=P(A)$ $(\forall\ A\in\mathscr{G})$. 显然，$P_*$为$(\Omega,\mathscr{G})$上的概率测度. 设$\zeta\in L'$，且$\zeta\geqslant 0$. 定义

$$\nu(A)=\int_A\zeta\mathrm{d}P\ (\forall\ A\in\mathscr{G}),$$

则 ν 为$(\Omega,\mathscr{G})$上的 P_*-连续测度，即$\nu\ll P_*$. 按定理 3.1.3，$\exists$ $(\mathrm{mod}P_*)$唯一且有限的非负$\mathscr{G}$-可测函数g，使得

$$\int_A\zeta\mathrm{d}P=\nu(A)=\int_A g\mathrm{d}P_*=\int_A g\mathrm{d}P\ (\forall\ A\in\mathscr{G}).$$

一般称g为ζ在条件$\mathscr{G}$下的条件期望.

定义3.2.1 设ζ为r.v.且$E\zeta$存在. 若定义在Ω上的可测映射Y(取值于$\overline{\mathrm{R}}$)，具有下列性质：

(i) Y是$\mathscr{G}$-可测的，即$Y^{-1}(B)\in\mathscr{G}$ $(\forall\ B\in\mathscr{B})$;

(ii) $\int_A Y\mathrm{d}P=\int_A\zeta\mathrm{d}P$ $(\forall\ A\in\mathscr{G})$;

则称Y为ζ在条件$\mathscr{G}$下的条件期望. 记为

$$Y=E(\zeta/\mathscr{G}).$$

例3.2.1 设$\mathscr{D}=(A_n,\ n\geqslant 1)$为$\Omega$的一个$\mathscr{F}$-集可数分割. $\mathscr{G}=\sigma(\mathscr{D})$，$\zeta\in L^1$. 问条件期望$Y=E(\zeta/\mathscr{G})$具有什么样的结构. 按可测映射的表现定理，有

$$Y=\sum_{n=1}^{\infty}a_n 1_{A_n}.$$

按定义 3.2.1，有

$$a_nP(A_n)=\int_{A_n}Y\mathrm{d}P=\int_{A_n}\zeta\mathrm{d}P\ (\forall\ n\geqslant 1).$$

若 $P(A_n)>0$，则 $a_n=\frac{1}{P(A_n)}\int_{A_n}\zeta\mathrm{d}P$；若 $P(A_n)=0$，则 a_n 可取任意值. 归纳起来，得

$$E(\zeta/\mathscr{G})=\sum_{n=1}^{\infty}a_n1_{A_n},\quad a_n=\frac{1}{P(A_n)}\int_{A_n}\zeta\mathrm{d}P\ \text{当}\ P(A_n)>0\ \text{时}.$$

(1) 假定 $E(\zeta/A_n)$ 表示 ζ 在事件 A_n 发生的条件下的条件期望，则

$$E(\zeta/A_n)=\frac{1}{P(A_n)}\int_{A_n}\zeta\mathrm{d}P=E(\zeta/\mathscr{G})(\omega)\ (\forall\ \omega\in A_n).$$

(2) 假如 $A_0\in\mathscr{G}$ 具有特点：$A_0\cap E=A_0$ 或 $\varnothing$ $(\forall\ E\in\mathscr{G})$，则称 A_0 为 $\mathscr{G}$ 中的一个孤立原子. 对此，有

$$E(\zeta/\mathscr{G})=E(\zeta/A_0)\ (\forall\ \omega\in A_0).$$

事实上，令 $Y=a1_{A_0}+E(\zeta/\mathscr{G})1_{A_0^c}$，则 Y 显然是 $\mathscr{G}$-可测的，且对 $\forall\ B\in\mathscr{G}$，有

$$\begin{aligned}\int_B Y\mathrm{d}P&=aP(B\cap A_0)+\int_{B\cap A_0^c}E(\zeta/\mathscr{G})\mathrm{d}P\\&=aP(BA_0)+\int_{BA_0^c}\zeta\mathrm{d}P.\end{aligned}$$

若 $BA_0=A_0$，且 $a=E(\zeta/A_0)$，则

$$\int_B Y\mathrm{d}P=aP(BA_0)+\int_{BA_0^c}\zeta\mathrm{d}P=\int_B\zeta\mathrm{d}P.$$

若 $BA_0=\varnothing$，则 $\int_B Y\mathrm{d}P=\int_{BA_0^c}\zeta\mathrm{d}P=\int_B\zeta\mathrm{d}P$.

$$\Rightarrow E(\zeta/\mathscr{G})=\begin{cases}E(\zeta/A_0), & \text{当}\ P(A_0)>0\ \text{时},\\ \text{任意值}, & \text{当}\ P(A_0)=0\ \text{时}.\end{cases}$$

在上例中，构作了一个符合定义 3.2.1 的 r.v. Y，那么是否还有其它的 Y 也符合定义 3.2.1 的要求呢？这就是所谓条件期望的唯一性问题. 另一个问题是，当 $E\zeta$ 存在时，$E(\zeta/\mathscr{G})$ 是否存在？假

如 $\zeta\in L^1$，条件期望的存在与唯一性问题可由定理 3.1.3 解决. 但若只有 $E\zeta$ 存在，比如 $E\zeta=\infty$，则定理 3.1.3 就不能用了. 因为由 $\nu(A)=\int_A\zeta\mathrm{d}P$ 所定义的测度 ν 不必 σ-有限. 例如，让 $\mathscr{D}$ 按例 3.2.1 定义，$P(A_n)>0\ (\forall\ n\geqslant1)$，$\zeta=\sum\limits_{n=1}^{\infty}\frac{1}{nP(A_n)}1_{A_n}$，则 $E\zeta=\sum\limits_n\frac{1}{n}=\infty$. 选 $\mathscr{G}=\{\varnothing,\Omega\}$，并定义

$$\nu(A)=\int_A\zeta\mathrm{d}P\quad(\forall\ A\in\mathscr{G}),$$

则 $\nu(\varnothing)=0$，$\nu(\Omega)=\infty$. 此测度虽有 $\nu\ll P$，但不具有 σ-有限性. 考虑到这种情形，就有必要推广定理 3.1.3.

引理 3.2.1 设 ν 为 $(\Omega,\mathscr{F})$ 上的 P-连续测度，则 $\exists\ B\in\mathscr{F}$，使 ν 在 $\mathscr{F}\cap B$ 上为 σ-有限测度，而对 $\forall\ A\in\mathscr{F}\cap B^c$，有

$$\nu(A)=\begin{cases}0, & P(A)=0,\\ \infty, & P(A)>0.\end{cases}$$

证 令

$$\mathscr{D}=\{D:\ D\in\mathscr{F},\ \text{且}\ \nu\ \text{在}\ \mathscr{F}\cap D\ \text{上为}\ \sigma\text{-有限测度}\},$$

$$\alpha=\sup\{P(D):\ D\in\mathscr{D}\}.$$

显然，$\mathscr{D}$ 是非空的. 在 $\mathscr{D}$ 中选子列 $(D_n,\ n\geqslant1)$，使得

$$\alpha=\lim_{n\to\infty}P(D_n).$$

让 $B=\bigcup\limits_n D_n$，则 $B\in\mathscr{D}$. 事实上，按 $\mathscr{D}$ 之定义，存在子列 $(D_n^k,\ k\geqslant1)\subset\mathscr{F}\cap D_n$，使 $\nu(D_n^k)<\infty\ (\forall\ k\geqslant1)$，且 $D_n=\bigcup\limits_n D_n^k$. 于是，

$$B=\bigcup_{n=1}^{\infty}\bigcup_{k=1}^{\infty}D_n^k=\bigcup_{n=1}^{\infty}\bigcup_{k=1}^{n}(D_n^k\cup D_k^n).$$

此式右边的每一项的 ν 测度都是有限的. 故 $B\in\mathscr{D}$. 此外，$P(B)=\alpha$. 事实上，

$$\alpha\geqslant P(B)\geqslant P(D_n)\Rightarrow P(B)\geqslant\lim_{n\to\infty}P(D_n)=\alpha\Rightarrow P(B)=\alpha.$$

现在，设 $A\in\mathscr{F}\cap B^c$. 若 $P(A)>0$，且 $\nu(A)<\infty$，则 $A\in\mathscr{D}$. $\Rightarrow\ \alpha\geqslant P(B\cup A)=P(B)+P(A)\geqslant\alpha\ \Rightarrow\ P(A)=0$. 这就出现矛

盾. 故 $P(A)>0 \Rightarrow \nu(A)=\infty$. 如果 $P(A)=0$,则由 $\nu \ll P$ 立即得: $\nu(A)=0$.

引理 3.2.2(R.-N. 定理的推广)　设 ν 为 $(\Omega;\mathscr{F})$ 上的 P-连续测度,则(mod$P$)存在唯一非负 $\mathscr{F}$-可测映射 f,使得 $\nu(A)=\int_A f\mathrm{d}P\ (\forall\ A\in\mathscr{F})$.

若 ν 是 σ-有限的,则 f(mod$P$)有限.

证　第二结论就是 R.-N. 定理. 下证第一结论.

设 B 为引理 3.2.1 中所确定的可测集. 将 ν, P 限制到 $\mathscr{F}\cap B$ 上. 记为 ν^*, P^*,且 ν^* 为 $\mathscr{F}\cap B$ 上的 σ-有限测度,则 $\nu^*\ll P^*$. 按定理 3.1.3,$\frac{\mathrm{d}\nu^*}{\mathrm{d}P^*}$ 在 B 上(modP^*)存在唯一且有限. 定义

$$\frac{\mathrm{d}\nu}{\mathrm{d}P}=\begin{cases}\dfrac{\mathrm{d}\nu^*}{\mathrm{d}P^*}, & \omega\in B,\\[2ex] \infty, & \omega\notin B.\end{cases}$$

则 $\frac{\mathrm{d}\nu}{\mathrm{d}P}$ 是 $\mathscr{F}$-可测的,且 $\nu(A)=\int_A\frac{\mathrm{d}\nu}{\mathrm{d}P}\mathrm{d}P\ (\forall\ A\in\mathscr{F})$.

事实上,对 $\forall\ A\in\mathscr{F}$,有

$$\nu(A)=\nu(AB)+\nu(AB^c),$$

其中 $\nu(AB)=\int_{AB}\frac{\mathrm{d}\nu^*}{\mathrm{d}P^*}\mathrm{d}P^*=\int_{AB}\frac{\mathrm{d}\nu}{\mathrm{d}P}\mathrm{d}P$.

$$\nu(AB^c)=\left.\begin{cases}\infty, & P(AB^c)>0,\\ 0, & P(AB^c)=0\end{cases}\right\}=\int_{AB^c}\frac{\mathrm{d}\nu}{\mathrm{d}P}\mathrm{d}P.$$

$$\Rightarrow \nu(A)=\int_A\frac{\mathrm{d}\nu}{\mathrm{d}P}\mathrm{d}P.$$

现在,令 $f=\frac{\mathrm{d}\nu}{\mathrm{d}P}$,则 f 即为所求. 为得此结论,仅需证明:(mod$P$)唯一. 假如 g 为另一满足要求的可测映射,则按定理 3.1.3,有 $f(\omega)=g(\omega)\ (\forall\ \omega\in B)$(mod$P$). 因此,$A=\{f\neq g\}\in\mathscr{F}\cap B^c$(mod$P$). 若 $P(A)>0$,则因 $f=\infty$ 于 A 上(mod$P$)必有 $\lambda>0$,使得 $P(\{g\leqslant\lambda\}\cap B^c)>0$.

$$\Rightarrow \nu(\{g\leqslant\lambda\}B^c)=\int_{\{g\leqslant\lambda\}B^c}g\mathrm{d}P\leqslant\lambda P(\{g\leqslant\lambda\}B^c)<\infty.$$

然而，按引理 3.2.1，$\nu(\{g\leqslant\lambda\}\cap B^c)=\infty$. 这就出现矛盾. 故

$$P(A)=0,\text{ 即 } f=g\ (\mathrm{mod}P).$$

（以后称$\frac{\mathrm{d}\nu}{\mathrm{d}P}$为广义 R.-N. 导数.）

定理 3.2.1 设$(\Omega,\mathscr{F},P)$为某概率空间. $\mathscr{G}$ 为 $\mathscr{F}$ 中的子 σ-代数. ζ 为 $\mathscr{F}$-可测函数(广义 r.v.)，且 $E\zeta$ 存在. 则(mod$P$)唯一存在某$\mathscr{G}$-可测函数 Y，使得

$$\int_A Y\mathrm{d}P=\int_A\zeta\mathrm{d}P\ (\forall\ A\in\mathscr{G}).$$

证 不妨设$\zeta\geqslant0$(mod$P$). 定义

$$\nu(A)=\int_A\zeta\mathrm{d}P\ (\forall\ A\in\mathscr{G}),$$

则 $\nu\ll P_*$，$P_*=P|_{\mathscr{G}}$. 按引理 3.2.2，广义 R.-N. 导数$\frac{\mathrm{d}\nu}{\mathrm{d}P_*}$(mod$P_*$)存在，唯一且$\mathscr{G}$-可测. 令$Y=\frac{\mathrm{d}\nu}{\mathrm{d}P_*}$，则

$$\int_A Y\mathrm{d}P=\int_A\frac{\mathrm{d}\nu}{\mathrm{d}P_*}\mathrm{d}P=\int_A\frac{\mathrm{d}\nu}{\mathrm{d}P_*}\mathrm{d}P_*=\nu(A)$$

$$=\int_A\zeta\mathrm{d}P.\ (\forall\ A\in\mathscr{G}).$$

推论 3.2.1 设$\mathscr{G}$ 为 $\mathscr{F}$ 中的子 σ-代数，$\zeta\in L'$. 若 $\sigma(\zeta)$独立于$\mathscr{G}$，则 $E(\zeta/\mathscr{G})=E\zeta$(mod$P$).

证 $\sigma(\zeta)$独立于 $\mathscr{G}$ 表明：对$\forall\ A\in\mathscr{G}$，有

$$\int_A E(\zeta/\mathscr{G})\mathrm{d}P=\int_A\zeta\mathrm{d}P=E(1_A\cdot\zeta)=P(A)\cdot E\zeta=\int_A E\zeta\mathrm{d}P$$

$$\Rightarrow E(\zeta/\mathscr{G})=E\zeta(\mathrm{mod}P).$$

推论 3.2.2 设 ζ 为$(\Omega,\mathscr{F})$上的 m-维随机矢量. Y 为$(\Omega,\mathscr{F})$上的 n-维随机矢量. ζ,Y 独立，f 为 $\mathrm{R}^{m+n}\to\mathrm{R}$ 的 Borel 函数，且 $Ef(\zeta,Y)$存在. 令

$$g(x)=\begin{cases}Ef(x,Y), & \text{当 } Ef(x,Y) \text{ 存在时},\\ 0, & \text{否则}.\end{cases}\quad(\forall\ x\in\mathrm{R}^m),$$

则 g 为 $\mathrm{R}^m\to\mathrm{R}$ 的 Borel 函数，且 $g(\zeta)=E(f(\zeta,Y)/\sigma(\zeta))$.

证 设$F_\zeta(x)$, $F_Y(y)$, $F_{\zeta Y}(x,y)$分别为ζ, Y, (ζ,Y)的分布函数. 对$\forall\ x\in\mathrm{R}^m$, $f(x,y)$为$y\in\mathrm{R}^n$上的Borel函数. 按Fubini定理,$g^{\pm}(x)$为$x\in\mathrm{R}^m$上的Borel函数. 这里

$$g^{\pm}(x)=Ef^{\pm}(x,Y)=\int_{\mathrm{R}^n}f^{\pm}(x,y)\mathrm{d}F_Y(y).$$

令$D=\{x\in\mathrm{R}^m: g^+(x)=\infty=g^-(x)\}$, $D^c=\mathrm{R}^m\backslash D$,则$D,D^c$均为$m$-维Borel集. 因此,$g(x)=(g^+(x)-g^-(x))1_D(x)$为Borel函数. 故$g(\zeta)$是$\sigma(\zeta)$-可测的. 按假设及Fubini定理,

$$\begin{aligned}Ef(\zeta,Y)&=\int_{\mathrm{R}^m\times\mathrm{R}^n}f(x,y)\mathrm{d}F_{\zeta Y}(x,y)\\&=\int_{\mathrm{R}^m}\int_{\mathrm{R}^n}f(x,y)\mathrm{d}F_Y(y)\mathrm{d}F_\zeta(x)\\&=\int_{\mathrm{R}^m}g(x)\mathrm{d}F_\zeta(x).\end{aligned}$$

按定理之条件,$Eg(\zeta)$存在. 按$\sigma(\zeta)$之定义,对$\forall\ A\in\sigma(\zeta)$, $\exists\ B\in\mathscr{B}^m$, $\ni A=\zeta^{-1}(B)$. 于是,

$$\begin{aligned}\int_A g^{\pm}(\zeta)\mathrm{d}P&=\int_B\int_{\mathrm{R}^n}f^{\pm}(x,y)\mathrm{d}F_Y(y)\mathrm{d}F_\zeta(x)\\&=\int_{B\times\mathrm{R}^n}f^{\pm}(x,y)\mathrm{d}F_{\zeta Y}(x,y)=\int_A f^{\pm}(\zeta,Y)\mathrm{d}P\end{aligned}$$

$$\Rightarrow\int_A g(\zeta)\mathrm{d}P=\int_A f(\zeta,Y)\mathrm{d}P\ (\forall\ A\in\sigma(\zeta)).$$

$$\Rightarrow g(\zeta)=E(f(\zeta,Y)/\sigma(\zeta)).\ (\mathrm{mod}P).$$

定理3.2.2 设$\zeta\in L^1$,$\mathscr{A}$为某集类,Y为$\sigma(\mathscr{A})$-可测的r.v.且$\int_A Y\mathrm{d}P=\int_A\zeta\mathrm{d}P\ (\forall\ A\in\mathscr{A})$. 若(i)$\mathscr{A}$为$\pi$类,且$\Omega\in\mathscr{A}$. 或(ii)$\mathscr{A}$为半环,且$\Omega$可表示成$\mathscr{A}$中可数个元素之并,则

$$Y=E(\zeta/\sigma(\mathscr{A}))(\mathrm{mod}P).$$

证 令$\mathscr{D}=\left\{A\in\sigma(\mathscr{A}): \int_A Y\mathrm{d}P=\int_A\zeta\mathrm{d}P\right\}$,则$\mathscr{A}\subseteq\mathscr{D}\subset\sigma(\mathscr{A})$. 若$\mathscr{A}$为$\pi$-类且$\Omega\in\mathscr{A}$,则"定理真"等价于"验证$\mathscr{D}$为$\lambda$-类". 显然,$\Omega\in\mathscr{D}$. 假定$A,B\in\mathscr{D}$,且$B\subset A$,则

$$\int_{A\setminus B} Y\mathrm{d}P = \int_A Y\mathrm{d}P - \int_B Y\mathrm{d}P = \int_{A\setminus B} \zeta\mathrm{d}P \Rightarrow A\setminus B \in \mathscr{D}.$$

现设$(A_n, n\geqslant1)\subset\mathscr{D}$，且$A_n\uparrow A$，则

$$\int_A \zeta\mathrm{d}P = E(1_A\cdot\zeta) = \lim_{n\to\infty}E(1_{A_n}\cdot\zeta) = \lim_{n\to\infty}E(1_{A_n}\cdot Y)$$

$$= E(1_A\cdot Y) = \int_A Y\mathrm{d}P.$$

这里用到 Lebesque 控制收敛定理. 故 $A\in\mathscr{D}$. 归纳起来便可得出 $\mathscr{D}$ 为 λ-类.

若 $\mathscr{A}$ 为半环，且 $\Omega=\sum_{n=1}^{\infty}A_n, A\in\mathscr{A}$，则类似于上述作法，不难验证 $\mathscr{D}$ 为 λ-类.

故定理在(i)和(ii)两种情形下真.

注 上述定理对 $E\zeta$ 存在的情形亦真. 对此，不妨设 $\zeta\geqslant0$. 令 $\zeta^c=\zeta 1_{\{\zeta\leqslant c\}}$，则定理对 ζ^c 成立. 然后，利用单调收敛定理，让 $c\uparrow\infty$，即可得这里所要的结论.

下面介绍一类常用的条件期望.

定义 3.2.2 设 ξ 为$(\Omega,\mathscr{F},P)$上的 r.v. η 为$(\Omega,\mathscr{F})\to(\Omega',\mathscr{F}')$的随机元. 令 $\mathscr{G}_\eta=\{A: A=\eta^{-1}(B), B\in\mathscr{F}'\}\subset\mathscr{F}$. 假如 $E\xi$ 存在，则称 $E(\xi/\mathscr{G}_\eta)$为 ξ 关于 η 的**条件期望**. 记为

$$E(\xi/\mathscr{G}_\eta) = E(\xi/\eta).$$

定理 3.2.3 设 ξ,η 为$(\Omega,\mathscr{F},P)$上的 r.v. 且 $E\xi$ 存在. 则$\exists$一个 Borel 函数 g，$\ni$ $g(\eta)=E(\xi/\eta)(\mathrm{mod}P)$，其中 g 由下式决定：

$$\int_B g\mathrm{d}P_\eta = \int_{\eta^{-1}(B)} \xi\mathrm{d}P \quad (\forall\ B\in\mathscr{B}).$$

这里 P_η 是由 η 诱导的$(\mathrm{R},\mathscr{B})$上的概率分布.

证 按第一章中所述的可测函数表现定理，$\exists$ Borel 函数 g，$\ni$ $g(\eta)=E(\xi/\eta)$. 剩下的问题是决定这个函数 g. 为此，不妨设 $\xi\geqslant0$. 定义 ν^*：

$$\nu^*(B) = \int_{\eta^{-1}(B)} \xi \mathrm{d}P \ (\forall\ B \in \mathscr{B}).$$

显然,$\nu^* \ll P_\eta$. 按引理 3.2.2,$\exists$ Borel 函数 $g = \frac{\mathrm{d}\nu^*}{\mathrm{d}P^\eta}$, 使得

$$\nu^*(B) = \int_B g(x) P_\eta(\mathrm{d}x) \ (\forall\ B \in \mathscr{B}).$$

令 $g^c(x) = g(x) 1_{\{g \leqslant c\}}(x)$,则

$$\int_B g^c(x) P_\eta(\mathrm{d}x) = \int_{\eta^{-1}(B)} g^c(\eta) \mathrm{d}P \ (\forall\ c > 0,\ B \in \mathscr{B}).$$

按单调收敛定理,得

$$\int_B g(x) P_\eta(\mathrm{d}x) = \int_{\eta^{-1}(B)} g(\eta) \mathrm{d}P$$

$$\Rightarrow \int_{\eta^{-1}(B)} g(\eta) \mathrm{d}P = \nu^*(B) = \int_{\eta^{-1}(B)} \xi \mathrm{d}P \ (\forall\ B \in \mathscr{B}).$$

$$\Rightarrow \int_B g \mathrm{d}P_\eta = \int_{\eta^{-1}(B)} \xi \mathrm{d}P \ (\forall\ B \in \mathscr{B}).$$

定理 3.2.4 设 ξ 为 $(\Omega, \mathscr{F}, P)$ 上的 r.v. 且 $E\xi$ 存在. η 为 $(\Omega, \mathscr{F}) \to (\mathrm{R}^T, \mathscr{B}^T)$ 的随机元,其中 $T = [a, b] \subset \mathrm{R}$,则

$\exists$ 一个 $(\mathrm{R}^T, \mathscr{B}^T) \to (\mathrm{R}, \mathscr{B})$ 的可测泛函 φ,使得

$$\varphi(\eta) = E(\xi/\eta) (\mathrm{mod} P).$$

此外,$\exists$ 一个可数集 $(t_k,\ k \geqslant 1) \subset T$ 及 $(\mathrm{R}^\infty, \mathscr{B}^\infty) \to (\mathrm{R}, \mathscr{B})$ 的可测映射 f,使得 $\varphi(\eta) = f(\eta_{t_1}, \eta_{t_2}, \cdots)$.

证 利用第一章中所述的可测函数的表现定理,立即可得此结论.

§ 3.3 条件期望的性质

设 $(\Omega, \mathscr{F}, P)$ 为某概率空间. $\mathscr{G}, \mathscr{G}_i$ 均表示 $\mathscr{F}$ 中的子 σ-代数. 定义算子 T、T_i:

$$T\zeta = E(\zeta/\mathscr{G}),\ T_i\zeta = E(\zeta/\mathscr{G}_i),\ (\zeta \in \mathscr{M}).$$

其中 $\mathscr{M} = \{\zeta:\ \zeta \text{ 是 } \mathscr{F}\text{-可测的,且 } E\zeta \text{ 存在}\}$.

本节将在等式或不等式之后省去说明“mod$P$”.

引理 3.3.1 算子 T 具有下列性质：

1. $T1=1$, $T(c\zeta)=cT\zeta$ $(c\in\mathrm{R},\ \zeta\in\mathscr{M})$；

2. $T(\zeta_1+\zeta_2)=T\zeta_1+T\zeta_2(\zeta_i\in\mathscr{M}$，且$(E|\zeta_1|)\wedge(E|\zeta_2|)<\infty)$；

3. $T\zeta=\zeta$ $(\zeta\in\mathscr{M}$，且 ζ 是$\mathscr{G}$-可测的)；

4. $T\zeta\geqslant 0$ $(0\leqslant\zeta\in\mathscr{M})$；

5. $T\zeta=E\zeta$ $(\zeta\in\mathscr{M}$，且与$\mathscr{G}$独立)；

6. $T_1(T_2\zeta)=T_2(T_1\zeta)$ $(\mathscr{G}_1\subset\mathscr{G}_2$ 或 $\mathscr{G}_2\subset\mathscr{G}_1,\zeta\in\mathscr{M})$；

7. $E\zeta=ET\zeta$ $(\zeta\in\mathscr{M})$.

证 每条性质都可由条件期望的定义直接证明.例如第 6 条之证如下.设 $\mathscr{G}_1\subset\mathscr{G}_2$，则按第 3 条，有 $E(E(\zeta|\mathscr{G}_1)|\mathscr{G}_2)=E(\zeta/\mathscr{G}_1)$. 对$\forall\ A\in\mathscr{G}_1$，自然有 $A\in\mathscr{G}_2$. 按定义 3.2.1，有

$$\int_A E(E(\zeta|\mathscr{G}_2)|\mathscr{G}_1)\mathrm{d}P=\int_A E(\zeta/\mathscr{G}_2)\mathrm{d}P=\int_A\zeta\mathrm{d}P$$
$$=\int_A E(\zeta/\mathscr{G}_1)\mathrm{d}P.$$

按条件期望的唯一性，立即可得所要之结论.

定理 3.3.1 设 Y 为$\mathscr{G}$-可测的 r.v. $\zeta\in\mathscr{M}$，且 $Y\cdot\zeta\in\mathscr{M}$，则 $T(Y\zeta)=Y\cdot T\zeta$.

证 不妨设 $\zeta\geqslant 0$, $Y\geqslant 0$. 定义 $\mathscr{F}$ 上的测度 ν,μ：

$$\nu(A)=\int_A Y\zeta\mathrm{d}P,\ \mu(A)=\int_A\zeta\mathrm{d}P\ (\forall\ A\in\mathscr{F}).$$

则 $\nu\ll P$, $\mu\ll P$. 令 $\nu^*=\nu|_{\mathscr{G}}$, $\mu^*=\mu|_{\mathscr{G}}$, $P^*=P|_{\mathscr{G}}$，则

$$\nu^*\ll P^*,\ \mu^*\ll P^*.$$

按定理 3.2.1，有

$$\frac{\mathrm{d}\nu^*}{\mathrm{d}P^*}=T(Y\cdot\zeta),\ \frac{\mathrm{d}\mu^*}{\mathrm{d}P^*}=T\zeta,\ \frac{\mathrm{d}\mu}{\mathrm{d}P}=\zeta.$$

按 R.-N. 导数的性质，对$\forall\ A\in\mathscr{G}$，有

$$\int_A Y\zeta\mathrm{d}P=\int_A Y\frac{\mathrm{d}\mu}{\mathrm{d}P}\mathrm{d}P=\int_A Y\mathrm{d}\mu=\int_A Y\mathrm{d}\mu^*=\int_A Y\frac{\mathrm{d}\mu^*}{\mathrm{d}P^*}\mathrm{d}P^*$$

$$=\int_A YT\zeta \mathrm{d}P^* = \int_A YT\zeta \mathrm{d}P.$$

$\Rightarrow YT\zeta \doteq E(Y\zeta/\mathscr{G}) = T(Y\cdot\zeta)$ (由唯一性).

定理 3.3.2(收敛性定理) 1. 单调收敛定理. 设 $Y\in L'$. $\zeta_n\in\mathscr{M}(n\geqslant 1)$. $Y\leqslant\zeta_n\uparrow\zeta$(a.s). 则 $\zeta\in\mathscr{M}$, 且 $T\zeta_n\uparrow T\zeta$(a.s).

2. Fatou 引理. 设 $Y\in L'$, $Y\leqslant\zeta_n\in\mathscr{M}(n\geqslant 1)$. 则

$$T(\underline{\lim}\zeta_n)\leqslant\underline{\lim}T\zeta_n \text{(a.s)}.$$

3. Lebesque 控制收敛定理. 设 $\zeta_n\in\mathscr{M}$, $|\zeta_n|\leqslant Y\in L'(n\geqslant 1)$, 且 $\zeta_n\xrightarrow{\text{a.s}}\zeta$. 则 $\zeta\in L'$, 且 $\lim\limits_{n\to\infty}T\zeta_n=T\zeta$(a.s).

4. 设 $\zeta_n\in\mathscr{M}$, 且 $\zeta_n\xrightarrow{P.}\zeta$. 若$(\zeta_n, n\geqslant 1)$一致可积,则 $\zeta\in L'$, 且 $T\zeta_n\xrightarrow{P.}T\zeta$. (注意:不必(a.s) 收敛).

证 $\boxed{\Rightarrow 1}$ 按期望的单调收敛定理,有

$$\int_A T\zeta_n\mathrm{d}P=\int_A\zeta_n\mathrm{d}P=E(1_A\zeta_n)\to E(1_A\zeta)=\int_A\zeta\mathrm{d}P$$

$$=\int_A T\zeta\mathrm{d}P\ (\forall\ A\in\mathscr{G}).$$

按假设,$(T\zeta_n, n\geqslant 1)$单调不降. 因此,$\lim\limits_{n\to\infty}T\zeta_n=z$(a.s)存在. 显然, z 是$\mathscr{G}$-可测的. 设

$$c_1=\{z>T\zeta\},\ c_2=\{z<T\zeta\},\ c=\{z\neq T\zeta\}=c_1\cup c_2.$$

则 $c, c_1, c_2\in\mathscr{G}$. 因此, $\int_{c_i}T\zeta_n\mathrm{d}P\to\int_{c_i}z\mathrm{d}P\ (i=1,2)$. 从而, $\int_{c_i}z\mathrm{d}P=\int_{c_i}T\zeta\mathrm{d}P$. 此式不能成立,除非 $P(c_i)=0\ (i=1,2)$

$\Rightarrow z=T\zeta$(a.s).

$\zeta\in\mathscr{M}$ 是显然的. 这就完成了结论 1 之证.

$\boxed{\Rightarrow 2}$ 此结论转化成结论 1 并不困难.

$\boxed{\Rightarrow 3}$ 利用结论 2 立即可推出此结论.

$\boxed{\Rightarrow 4}$ 设 $\delta>0$, 令 $\zeta_n^c=\zeta_n 1_{\{|\zeta_n|\leqslant c\}}$, $\zeta^c=\zeta 1_{\{|\zeta|\leqslant c\}}$, 则

$$
\begin{aligned}
&P(|T\zeta_n - T\zeta| > 3\delta) \\
\leqslant & P(|T\zeta_n - T\zeta_n^c| > \delta) + P(|T\zeta_n^c - T\zeta^c| > \delta) \\
& + P(|T\zeta - T\zeta^c| > \delta).
\end{aligned}
$$

现在,估计此式右边的三项

$$
P(|T(\zeta_n - \zeta_n^c)| > \delta) \leqslant \frac{1}{\delta} E(T|\zeta_n - \zeta_n^c|)
$$

$$
= \frac{1}{\delta} E|\zeta_n - \zeta_n^c| \leqslant \frac{1}{\delta} \sup_n E|\zeta_n - \zeta_n^c| \xrightarrow[c \to \infty]{} 0;
$$

$$
P(|T(\zeta - \zeta^c)| > \delta) \leqslant \frac{1}{\delta} E|\zeta - \zeta^c| \xrightarrow[c \to \infty]{} 0;
$$

$$
P(|T(\zeta_n^c - \zeta^c)| > \delta) \leqslant \frac{1}{\delta} E|\zeta_n^c - \zeta^c|
$$

$$
\leqslant \frac{\varepsilon}{\delta} + \frac{2c}{\delta} P(|\zeta_n^c - \zeta^c| > \varepsilon) \xrightarrow[c \to \infty]{} \frac{\varepsilon}{\delta} \quad (\forall\, \varepsilon > 0).
$$

这里第一估式用到 ζ 的一致可积性,第二估式用到 $\zeta \in L'$,第三估式用到 $\zeta_n \xrightarrow{P.} \zeta$.

$$
\Rightarrow P(|T\zeta_n - T\zeta| > 3\delta) \xrightarrow[n \to \infty]{} 0 \quad (\forall\, \delta > 0).
$$

注 (i) 在结论 1,2 中,下控条件不能减弱.这完全类似于期望的情形.关于期望的 Lebesque 有界收敛定理的条件“$\zeta_n \xrightarrow{a.s} \zeta$”可以改成条件“$\zeta_n \xrightarrow{P.} \zeta$”.

(ii) 结论 3 中的条件“$\zeta_n \xrightarrow{a.s} \zeta$”不能改成“$\zeta_n \xrightarrow{P.} \zeta$”.这是因为若取 $\mathscr{G} = \mathscr{F}$,则 $T\zeta_n = \zeta_n$,$T\zeta = \zeta$,而 $\zeta_n \xrightarrow{P.} \zeta$ 并不意味着“$\zeta_n \xrightarrow{a.s} \zeta$”.此结论中的控制条件“$|\zeta_n| \leqslant Y \in L'$”也不能改成“$(\zeta_n)$ 一致可积”.有反例可证.

推论 3.3.1 设 $\zeta_n \in \mathscr{M}$,且 $\sup_n |\zeta_n| \in L'$. $\zeta_n \xrightarrow{a.s} \zeta$,则 $\zeta \in \mathscr{M}$,且 $T\zeta_n \xrightarrow{a.s} T\zeta$.

定理 3.3.3(基本不等式)

1. Jensen 不等式.设 g 为 R 上的下凸 Borel 函数,则 $g(T\zeta)$

$\leqslant Tg(\zeta)$ (a. s) ($\forall\ \zeta\in\mathscr{M}$, $T\zeta$ 为 r. v.);

2. Lyapunov 不等式. $(T|\zeta|^s)^{\frac{1}{s}}\leqslant(T|\zeta|^t)^{\frac{1}{t}}$ ($\zeta\in\mathscr{M}$, $0<s\leqslant t$);

3. Hölder 不等式. $T|\zeta\cdot\eta|\leqslant(T|\zeta|^p)^{\frac{1}{p}}(T|\eta|^q)^{\frac{1}{q}}$ ($\zeta,\eta\in\mathscr{M}$, $\frac{1}{p}+\frac{1}{q}=1$, $1<p,q<\infty$);

4. Minkowski 不等式

$$(T|\xi+\eta|^p)^{\frac{1}{p}}\leqslant(T|\xi|^p)^{\frac{1}{p}}+(T|\eta|^p)^{\frac{1}{p}}$$
$$(\xi,\eta\in\mathscr{M},\ 1\leqslant p<\infty);$$

5. $|T\zeta|^p\leqslant T|\zeta|^p$ ($\zeta\in\mathscr{M}$, $1\leqslant p<\infty$).

证 各不等式之证可转化成期望的相应情形.

§3.4 条件概率

设$(\Omega,\mathscr{F},P)$为概率空间. $\mathscr{G}$ 为 $\mathscr{F}$ 中的子 σ-代数. 定义 $\Omega\times\mathscr{F}$ 上的二元函数 $h(\cdot,\cdot)$:

$$h(\omega,A_0)=E(1_{A_0}/\mathscr{G})\quad((\omega,A_0)\in\Omega\times\mathscr{F}).$$

则 h 具有下列特性:

(i) 对$\forall\ \omega\in\Omega$, h 以概率 1 为 $\mathscr{F}$ 上的概率测度;

(ii) 对$\forall\ A_0\in\mathscr{F}$, h 是 $\mathscr{G}$-可测的;

(iii) $\int_A h(\omega,A_0)\mathrm{d}P=P(A\cap A_0)$ ($\forall\ A\in\mathscr{G}$, $A_0\in\mathscr{F}$).

定义 3.4.1 若定义在 $\Omega\times\mathscr{F}$ 上的二元函数 h 满足条件(i)~(iii),则称 h 为关于 $\mathscr{G}$ 的正则函数. 对每个 $A_0\in\mathscr{F}$,称正则函数 $h(\omega,A_0)$为 A_0 在条件 $\mathscr{G}$ 下的**正则条件概率**. 记为 $h(\omega,A_0)=E(1_{A_0}/\mathscr{G})$.

注 1 正则条件概率就是一种特殊的条件期望. 但若 h 仅满足条件(i)和(ii),则不必有这种特殊关系. 例如,设 $\mathscr{G}=\mathscr{F}$,G 为 $\mathscr{F}$ 上的概率测度. 令

$$h(\omega,A_0)=G(A_0).\ 其中\ A_0\in\mathscr{F},\ \omega\in\Omega.$$

则 $\int_A h(\omega,A_0)\mathrm{d}P=P(A)\cdot G(A_0)=P(AA_0)\ (\forall\ A\in\mathscr{F})$ 一般不成立. 由此可见,h 不必满足第(iii)条,但(i)和(ii)满足.

注 2 从条件期望的定义中可以看出:条件(i)蕴含在条件(ii)和(iii)中.

注 3 二元函数 h 要求在 $\Omega\times\mathscr{F}$ 上具有正则性. 这对许多实际问题显得有点苛刻. 例如,往往只要求 h 在 $\Omega\times\mathscr{F}'$ 上具有正则性就够了,其中 $\mathscr{F}'$ 为 $\mathscr{F}$ 中的某个子 σ-代数. 比如:$\mathscr{F}'=\sigma(\zeta)$. 这就引申出如下概念.

定义 3.4.2 设 ζ 为$(\Omega,\mathscr{F},P)$上的 r.v. 若定义在 $\Omega\times\mathscr{B}$ 的二元函数 Q 满足下列要求:

(i) 对$\forall\ \omega\in\Omega$, $Q(\omega,B)$以概率 1 为 $\mathscr{B}$ 上的概率分布;

(ii) 对$\forall\ B\in\mathscr{B}$, $Q(\omega,B)$是$\mathscr{G}$-可测的;

(iii) $\int_A Q(\omega,B)\mathrm{d}P=P(A\cap\zeta^{-1}(B))\ (\forall\ A\in\mathscr{G},\ B\in\mathscr{B})$.

则称 Q 为 $\Omega\times\mathscr{B}$ 上的正则函数.

注 4 设 $\mathscr{F}'=\sigma(\zeta)$, $P^*=P|\mathscr{F}'$, $h^*(\omega,A_0^*)=h(\omega,\zeta^{-1}(B_0))=Q(\omega,B_0)\ (A_0^*=\zeta^{-1}(B_0))$. 若 Q 为 $\Omega\times\mathscr{B}$ 上的正则函数,则 h^* 为 $\Omega\times\mathscr{F}'$ 上的正则函数.

定理 3.4.1(条件期望与条件概率的相互表示) 设 ξ 为r.v.,且 $E\xi$ 存在. $h(\omega,A)=E(1_A/\mathscr{G})$,则

$$E(\xi/\mathscr{G})=\int_\Omega \xi(\tilde{\omega})h(\omega,\mathrm{d}\tilde{\omega}).\ \ (\text{a.s}).$$

证 不妨设 $\xi\geqslant0$. 若 $\xi=1_A$. 则按定义 4.1,公式成立. 因此,当 ξ 为简单变量时,公式成立. 对于 r.v. $\xi\geqslant0$,$\exists$ 简单 r.v. 列$(\xi_n,\ n\geqslant1)$,使得 $\xi_n\geqslant0$, $\xi_n\uparrow\xi$ (a.s). 按单调收敛定理,即得公式一般成立.

定义 3.4.3 设 ξ 为$(\Omega,\mathscr{F},P)$上的 r.v., $\mathscr{G}$ 为 $\mathscr{F}$ 中的子 σ-代数. 若定义在 $\Omega\times\mathrm{R}$ 上的二元函数 F 具有下列性质:

(i) 对$\forall\ \omega\in\Omega,F(\omega,x)$以概率 1 为 R 上的分布函数；

(ii) 对$\forall\ x\in\mathrm{R},F(\omega,x)$是$\mathscr{G}$-可测的；

(iii) $\int_A F(\omega,x)P(\mathrm{d}\omega)=P(A\cdot\xi^{-1}((-\infty,x]))\ (\forall\ A\in\mathscr{G},\ x\in\mathrm{R})$.

则称 F 为 ξ 关于 $\mathscr{G}$ 的**正则条件分布函数**.

定理 3.4.2 正则条件分布 $F(\mathrm{mod}P)$存在且唯一.

证 设 Q_ξ 为 $\Omega\times\mathscr{B}$ 上关于 $\mathscr{G}$ 的正则条件概率，其中 Q_ξ 是在定义 3.4.2 中将 ζ 改成 ξ 所得到的. 由条件期望的存在与唯一性得到 Q_ξ 的存在与唯一性. 令 $F(\omega,x)=Q_\xi(\omega,(-\infty,x])$，则 F 为 ξ 关于 $\mathscr{G}$ 的正则条件分布函数. 按测度扩张的唯一性知 F 是唯一被决定的.

定义 3.4.4 若可测空间$(E,\mathscr{E})$等价于在某个 Borel 集 $G\subset\mathrm{R}$ 上产生的 Borel 可测空间$(G,\mathscr{B}_G)$，其中 $\mathscr{B}_G=\mathscr{B}\cap G$，即$\exists$ 一个 1-1 对应的映射 φ: $(E,\mathscr{E})\to(G,\mathscr{B}_G)$，于

(i) $\varphi(E)=G,\ \varphi^{-1}(G)=E$；

(ii) φ是$\mathscr{E}\backslash\mathscr{B}_G$-可测的；

(iii) φ^{-1}是 $\mathscr{B}_G\backslash\mathscr{E}$ 可测的；

则称$(E,\mathscr{E})$为一个 **Borel 空间**.

定理 3.4.3 设 ξ 是$(\Omega,\mathscr{F})\to(E,\mathscr{E})$的随机元. $(E,\mathscr{E})$为 Borel 空间. 则 ξ 关于 $\mathscr{G}$ 的正则条件概率存在且唯一，其中 $\mathscr{G}$ 为 $\mathscr{F}$ 中的子 σ-代数.

证 设 φ 为定义 3.4.4 中的映射，则 $\varphi(\xi(\omega))=\varphi\circ\xi(\omega)$为取值于$G$ 的 r.v. 相应于 $\varphi\circ\xi$ 和 $\mathscr{G}$ 的正则函数 $Q(\omega,B)$存在且唯一，其中$(\omega,B)\in\Omega\times\mathscr{B}_G$. 令

$$\tilde{Q}(\omega,A)=Q(\omega,\varphi(A))\ (\forall\ A\in\varepsilon).$$

则 $\tilde{Q}$在 $\Omega\times\mathscr{E}$ 上有定义，且

$$\begin{aligned}\tilde{Q}(\omega,A)&=Q(\omega,\varphi(A))=P(\varphi\circ\xi\in\varphi(A)/\mathscr{G})\\&=P(\xi\in A/\mathscr{G}).\end{aligned}$$

最后交待某些特殊记号，设ξ,η为$(\Omega,\mathscr{F},P)$上的r.v.常用的简化符号有

$$P(A/\xi)=P(A/\sigma(\xi)),\quad P(A/\xi=x)=P(A/\xi)|_{\xi=x},$$
$$E(\eta/\xi)=E(\eta/\sigma(\xi)),\quad E(\eta/\xi=x)=E(\eta/\xi)|_{\xi=x}.$$

这些符号有其直观性.但要特别注意，这些符号的合理性在于可测函数的表现定理.像$P(A/\xi=x)$一般不能解释为"在事件$\{\xi=x\}$下的条件概率".本书将尽量不采用这样的符号，但$P(A/\xi)$无论怎样解释都是合理的.

§3.5 因子分解

设$(P_\theta)_{\theta\in\Theta}$为$(\Omega,\mathscr{F})$上的概率测度族，其中$\Theta$为某个参数集.在统计学中有一个很重要的问题就是从观测ω推断参数θ，即从概率测度族(P_θ)中推断概率P_θ.

设$P_\theta(A|\mathscr{G})$和$E_\theta(X|\mathscr{G})$分别表示关于定义在$(\Omega,\mathscr{F})$上的概率测度P_θ的条件概率和条件期望，其中$\mathscr{G}$为$\mathscr{F}$中的子σ-代数.

定义 3.5.1 若$\exists$ 定义于$\mathscr{F}\times\Omega$上的函数p，使得对$\forall\ A\times\omega\in\mathscr{F}\times\Omega$和$\forall\ \theta\in\Theta$，有

$$p(A,\omega)=P_\theta(A/\mathscr{G})\ (\mathrm{mod}P_\theta),$$

则说$\mathscr{G}$关于族(P_θ)是充足的.

定义 3.5.2 若T为r.v.或随机矢量，且$\sigma(T)$为关于族(P_θ)的充足子σ-代数，则称T为关于族(P_θ)的**充足统计**.

注 充足子σ-代数$\mathscr{G}$这一概念表达如下含义：即使不知道$\mathscr{F}$的不含在$\mathscr{G}$中元素之信息，也不会影响关于θ的推断.

例 3.5.1 设$(\Omega,\mathscr{F})=(R^k,\mathscr{B}^k)$，$\Theta=\{\theta:\ \theta>0\}$，$\lambda_k$为$k$-维Lebesque测度，$P_\theta$关于$\lambda_k$的R.-N.导数$f_\theta(x)=\dfrac{\mathrm{d}P_\theta}{\mathrm{d}\lambda_k}(x)$具有如下结构：

$$f_\theta(x)=f_\theta(x_1,\cdots,x_k)=\begin{cases}\theta^{-k}, & \text{当 } 0\leqslant x_i\leqslant\theta,\ 1\leqslant i\leqslant k \text{ 时},\\ 0, & \text{否则}.\end{cases}$$

设 X_i 为 $\mathrm{R}^k \to \mathrm{R}$ 的可测函数，且 $X_i(x)=x_i(\forall\ x\in\mathrm{R}^k)$. 则随机元族$(X_i)_{1\leqslant i\leqslant k}$具有下列特性：

(1) (X_i)关于(P_θ)是独立的.

事实上，设 $B_i=[0,b_i]$，其中 $0\leqslant b_i\leqslant\theta$，则

$$P_\theta(X_i\in B_i,\ 1\leqslant i\leqslant k)=\int_0^{b_1}\cdots\int_0^{b_k}\theta^{-k}\mathrm{d}\lambda_k=\prod_{i=1}^k\left(\frac{b_i}{\theta}\right),$$

$$P_\theta(X_i\in B_i)=\frac{b_i}{\theta}$$

$$\Rightarrow P_\theta(X_i\in B_i,\ 1\leqslant i\leqslant k)=\prod_{i=1}^k\left(\frac{b_i}{\theta}\right)=\prod_{i=1}^k P_\theta(X_i\in B_i).$$

不难证明：此式对任意的 Borel 集族$(B_i)_{1\leqslant i\leqslant k}$是成立的.

(2) (X_i)关于(P_θ)的分布是均匀的.

事实上，X_1 关于 P_θ 的分布为 $F_\theta(x)=P_\theta(X_1\leqslant x_1)$. 因此，

$$F_\theta(x_1)=\int_{\mathrm{R}^{k-1}}\lambda_{k-1}(\mathrm{d}\omega^{k-1})\int_{-\infty}^{x_1}f_\theta(\omega_1,\ \omega^{k-1})\lambda_1(\mathrm{d}\omega_1)$$

$$=\begin{cases}0, & x_1\leqslant 0,\\ \dfrac{x_1}{\theta}, & 0<x_1\leqslant\theta,\quad \forall\ x_1\in\mathrm{R}.\\ 1, & \theta<x_1.\end{cases}$$

按独立性，(X_i)关于 P_θ 的联合分布函数 $\widetilde{F}_\theta(x)$为

$$\widetilde{F}_\theta(x)=\prod_{i=1}^k F_\theta(x_i),\quad \forall\ x=(x_1,x_2,\cdots,x_k)\in\mathrm{R}^k.$$

(3) 设 $T(\omega)=\max(X_i(\omega):\ 1\leqslant i\leqslant k)$，则 T 具有下列特性：

(i) T 的分布函数 $F_\theta^{(T)}(t)=F_\theta^k(t)\ (t\in\mathrm{R})$；

(ii) $E_\theta(X_i/T)=\dfrac{k+1}{2k}T\ (\mathrm{mod}P)$；

(iii) $\sigma(T)$关于(P_θ)是充足的.

以下分三步来证明这个结论.

首先证(i).

$$F_\theta^T(t)=P_\theta(T\leqslant t)=P_\theta(\{X_i\leqslant t\},\ 1\leqslant i\leqslant k)$$

$$=\prod_{i=1}^{k} P_\theta(X_i \leqslant t) = F_\theta^k(t).$$

其次论证(ii). 为此，考虑下列关系式:

$$\int_{\{T\leqslant t\}} E_\theta(X_i|T)\mathrm{d}P_\theta = \int_{\{T\leqslant t\}} X_i \mathrm{d}P_\theta,$$

$$\{T \leqslant t\} = \prod_{i=1}^{k} \{X_i \leqslant t\},$$

$$\int_{\{T\leqslant t\}} X_i \mathrm{d}P_\theta = \prod_{l\neq i} P_\theta(X_l \leqslant t) \cdot \int_{\{X_i\leqslant t\}} X_i \mathrm{d}P_\theta$$

$$=F_\theta^{k-1}(t) \cdot \int_{\{X_i\leqslant t\}} X_i \mathrm{d}P_\theta$$

$$=F_\theta^{k-1}(t) \cdot \int_{-\infty}^{t} u \mathrm{d}F_\theta(u).$$

$$\int_{\{T\leqslant t\}} T\mathrm{d}P_\theta = \int_{-\infty}^{t} u\mathrm{d}F_\theta^k(u) = \frac{2k}{k+1}\int_{\{T\leqslant t\}} X_i \mathrm{d}P_\theta$$

$$\Rightarrow \int_{\{T\leqslant t\}} E_\theta(X_i|T)\mathrm{d}P_\theta = \frac{k+1}{2k}\int_{\{T\leqslant t\}} T\mathrm{d}P_\theta \quad (\forall\, t \in R).$$

故 $E_\theta(X_i|T) = \frac{k+1}{2k} T \ (\mathrm{mod} P_\theta)$.

最后证明(iii).

显然，$\frac{\mathrm{d}P_\theta}{\mathrm{d}\lambda_k}(x) = \frac{1}{\theta^k} 1_{(0,\theta)^k}(x)$，其中$(0,\theta)^k = \mathop{\times}\limits_{1}^{k} (0,\theta)$.

因此，$\frac{\mathrm{d}P_\theta}{\mathrm{d}\lambda_k}$是$\sigma(T)$-可测的. 构作概率测度

$$P(A) = \sum_{n=1}^{\infty} 2^{-n} P_{\theta_n}(A) \ (\forall\ A \in \mathscr{F}),\ 其中\ \theta_n = n.$$

令 $f = \sum_{n=1}^{\infty} 2^{-n} \frac{\mathrm{d}P_{\theta_n}}{\mathrm{d}\lambda_k}$，则 f 是$\sigma(T)$-可测的，且

$$P(f=0) = 0.$$

于是，$\frac{\mathrm{d}P_\theta}{\mathrm{d}P} = \frac{1}{f} \frac{\mathrm{d}P_\theta}{\mathrm{d}\lambda_k}$ $(\mathrm{mod}P)$. 此式表明：$\frac{\mathrm{d}P_\theta}{\mathrm{d}P}$是$\sigma(T)$-可测的. 由此即知：对$\forall\ B \in \sigma(T)$，有

$$\int_B E_\theta(1_A/\sigma(T))\mathrm{d}P_\theta = \int_B 1_A \mathrm{d}P_\theta = \int_B 1_A \frac{\mathrm{d}P_\theta}{\mathrm{d}P}\mathrm{d}P$$

$$= \int_B \frac{dP_\theta}{dP} E(1_A/\sigma(T)) dP = \int_B E(1_A/\sigma(T)) dP_\theta$$

$$\Rightarrow E(1_A/\sigma(T)) = E_\theta(1_A/\sigma(T)) \ (\mathrm{mod} P_\theta).$$

按定义 3.5.1,$\sigma(T)$关于(P_θ)是充足的.

现在介绍因子分解定理.

引理 3.5.1 设$(P_\theta)_{\theta\in\Theta}$受测度 μ 的控制,即 $P_\theta \ll \mu (\forall\ \theta \in \Theta)$,且对$\forall\ \theta\in\Theta$,$P_\theta$ 关于 μ 具有密度

$$f_\theta = \frac{dP_\theta}{d\mu} = h \cdot g_\theta,$$

其中 g_θ 是$\mathscr{G}$-可测的,h 是非负 μ-可积的,则$\mathscr{G}$ 关于(P_θ)是充足的.

证 令 $\alpha = \int_\Omega h d\mu$, $f_\theta = \left(\frac{1}{\alpha} h\right) \cdot (\alpha \cdot g_\theta)$,则以 $\alpha \cdot g_\theta$ 代 g_θ,$\frac{h}{\alpha}$ 代 h 之后 f_θ 的结构不变. 因此,不妨设 $\int_\Omega h d\mu = 1$.

定义$\mathscr{F}$ 上的集函数 P: $P(A) = \int_A h d\mu \ (\forall\ A \in \mathscr{F})$,则 P 为$(\Omega,\mathscr{F})$上的概率测度,且对$\forall\ A\in\mathscr{F}$,有

$$P_\theta(A) = \int_A f_\theta d\mu = \int_A g_\theta \cdot h d\mu = \int_A g_\theta dP.$$

对$\forall\ G\in\mathscr{G}$,有

$$\int_G P(A/\mathscr{G}) dP_\theta = \int_G P(A/\mathscr{G}) \cdot g_\theta dP$$

$$= \int_G 1_A \cdot g_\theta dP = \int_G 1_A dP_\theta$$

$$\Rightarrow P(A/\mathscr{G}) = P_\theta(A/\mathscr{G}) \ (\mathrm{mod} P_\theta).$$

按定义 3.5.1,$\mathscr{G}$ 关于(P_θ)是充足的.

引理 3.5.2 设$(P_\theta)_{\theta\in\Theta}$同时受 μ 和 P_{θ_0} 的控制,其中 θ_0 为 Θ 中的某个元素. 若$\mathscr{G}$ 关于(P_θ)是充足的,则每个 P_θ 关于 μ 具有密度

$$f_\theta = h \cdot g_\theta$$

其中 g_θ 是$\mathscr{G}$-可测的,h 是非负 μ-可积的.

证　令 $q_\theta=\frac{\mathrm{d}P_\theta}{\mathrm{d}P_{\theta_0}}$，则对 $\forall\ A\in\mathscr{F}$，有

$$\begin{aligned}\int_A E_{\theta_0}(q_\theta/\mathscr{G})\mathrm{d}P_{\theta_0} &=\int 1_A\cdot E_{\theta_0}(q_\theta/\mathscr{G})\mathrm{d}P_{\theta_0}\\ &=\int E_{\theta_0}(1_A/\mathscr{G})\cdot q_\theta \mathrm{d}P_{\theta_0}\\ &=\int p(A,\omega)\cdot q_\theta(\omega)P_{\theta_0}(\mathrm{d}\omega)\ (\text{按假设条件})\\ &=\int p(A,\omega)\cdot P_\theta(\mathrm{d}\omega)=\int P_\theta(A/\mathscr{G})\mathrm{d}P_\theta\\ &=P_\theta(A)=\int_A f_\theta \mathrm{d}\mu\end{aligned}$$

$$\Rightarrow\int_A f_\theta\mathrm{d}\mu=\int_A E_{\theta_0}(q_\theta/\mathscr{G})\cdot\frac{\mathrm{d}P_{\theta_0}}{\mathrm{d}\mu}\mathrm{d}\mu\ (\forall\ A\in\mathscr{F})$$

$$\Rightarrow f_\theta=h\cdot g_\theta(\bmod\mu).$$

其中 $h=\frac{\mathrm{d}P_{\theta_0}}{\mathrm{d}\mu}$（$\mu$-可积），$g_\theta=E_{\theta_0}(q_\theta/\mathscr{G})$（$\mathscr{G}$-可测）.

引理 3.5.3　若 $(P_\theta)_{\theta\in\Theta}$ 受一个 σ-有限测度 μ 的控制，则它等价于它的某个可数子族 $(P_{\theta_n})_{n\geqslant1}$，即 $(P_\theta)_{\theta\in\Theta}$ 与 $(P_{\theta_n})_{n\geqslant1}$ 相互控制.

证　μ 的 σ-有限性表明：$\exists\ \Omega$ 的 $\mathscr{F}$-集分割 $(A_n)_{n\geqslant1}$，$\ni 0<\mu(A_n)<\infty$（$\forall\ n\geqslant1$）. 定义测度

$$P(A)=\sum_{n=1}^{\infty}2^{-n}\frac{\mu(A\cdot A_n)}{\mu(A_n)}\ (\forall\ A\in\mathscr{F}).$$

则 P 为 $\mathscr{F}$ 上的概率测度，且 $P\sim\mu$，即 $P\ll\mu\ll P$. 由此知，(P_θ) 受 P 的控制. 基于这一考虑，不妨设 μ 为有限测度.

令 $f_\theta=\frac{\mathrm{d}P_\theta}{\mathrm{d}\mu}$，$S_\theta=\{\omega:\ f_\theta(\omega)>0\}$，则

$$P_\theta(A)=\int_A f_\theta\mathrm{d}\mu=\int_{A\cdot S_\theta}f_\theta\mathrm{d}\mu=P_\theta(A\cdot S_\theta),$$

$$P_\theta(A)=0\Longleftrightarrow\mu(A\cdot S_\theta)=0.$$

这表明：S_θ 为 P_θ 的支撑集，且 $P_\theta\sim\mu$ 于 $\mathscr{F}\cap S_\theta$ 上.

如果 $B\in\mathscr{F}$，且 $\exists\ \theta\in\Theta$，$\ni B\subset S_\theta$，则称 B 为一个核，称有限

或可数个核之并为一个链. 在此约定下,令

$$\alpha = \sup\{\mu(C): C \text{ 为链}\},$$

则 α 是有限的,且 $\exists$ 某链 C, $\ni \mu(C)=\alpha$. 假定链 C 具有表示 $C=\bigcup_n B_n$,其中 B_n 为核,且 $B_n \subset S_{\theta_n}$.

现在,我们要证明:$(P_\theta)_{\theta\in\Theta}$ 受 $(P_{\theta_n})_{n\geqslant1}$ 的控制,这里 $(\theta_n)_{n\geqslant1}$ 为 Θ 中的可数子族. 从而,它们等价. 设

$$P_{\theta_n}(A) = 0 \ (\forall\ n \geqslant 1),$$

则 $\mu(A\cdot S_{\theta_n})=0$ $(\forall\ n\geqslant1)$. 从而,

$$\mu(A\cdot C) = \mu(\bigcup_n A\cdot B_n) \leqslant \mu(\bigcup_n A\cdot S_{\theta_n})$$

$$\leqslant \sum_n \mu(A\cdot S_{\theta_n}) = 0.$$

$$\Rightarrow P_\theta(A\cdot C) = 0\ (\forall\ \theta\in\Theta)\ (\text{按受控假设}).$$

为证 $P_\theta(A)=0$ $(\forall\ \theta\in\Theta)$,仅需证明:$P_\theta(A\backslash C)=0$ $(\forall\ \theta\in\Theta)$. 假如相反,即 $P_\theta(A\backslash C)>0$,则

$$P_\theta((A\backslash C)\cdot S_\theta) = P_\theta(A\backslash C) > 0$$

$$\Rightarrow \mu((A\backslash C)\cdot S_\theta) > 0.$$

注意,$(A\backslash C)\cdot S_\theta$ 是一个与 C 不相联的核. 因此,

$$\alpha = \mu(C) \geqslant \mu(C\cup (A\backslash C)\cdot S_\theta) = \alpha + \mu((A\backslash C)\cdot S_\theta) > \alpha.$$

这个矛盾的原因在于 $P_\theta(A\backslash C)>0$,故 $P_\theta(A\backslash C)=0$.

注 "$(P_\theta)_{\theta\in\Theta}$ 受 $(P_{\theta_n})_{n\geqslant1}$ 的控制"等价于"当 $P_{\theta_n}(A)=0$ $(\forall\ n\geqslant1)$ 时,$P_\theta(A)=0$ $(\forall\ \theta\in\Theta)$".

定理 3.5.1 设 $(P_\theta)_{\theta\in\Theta}$ 受一个 σ-有限度 μ 的控制. 则 $\mathscr{G}$ 关于 (P_θ) 是充足的充分必要条件为

$$f_\theta = \frac{\mathrm{d}P_\theta}{\mathrm{d}\mu} = h\cdot g_\theta,$$

其中 g_θ 是 $\mathscr{G}$-可测的,h 是非负 μ-可积的.

证 充分性. 假定分解式成立. 按引理 3.5.3,可选出子族 $(P_{\theta_n})_{n\geqslant1}$,使得:

$$(P_{\theta_n})_{n\geqslant 1} \sim (P_\theta)_{\theta\in\Theta}.$$

定义 $\mathscr{F}$ 上的概率 P：

$$P(A) = \sum_{n=1}^{\infty} 2^{-n} P_{\theta_n}(A) \ (\forall\ A \in \mathscr{F}).$$

则 $P \sim (P_{\theta_n}) \sim (P_\theta)$. 从而，$P \ll \mu$，且

$$\frac{\mathrm{d}P}{\mathrm{d}\mu} = \sum_{n=1}^{\infty} 2^{-n} \frac{\mathrm{d}P_{\theta_n}}{\mathrm{d}\mu} = \sum_{n=1}^{\infty} 2^{-n} g_{\theta_n} \cdot h = f \cdot h$$

其中 $f = \sum_{n=1}^{\infty} 2^{-n} g_{\theta_n}$ 是 $\mathscr{G}$-可测的. 易证：$P(f=0)=0$，$P_\theta(f=0)=0$ $(\forall\ \theta\in\Theta)$. 于是，$r_\theta = \frac{1}{f} g_\theta$ 是 $\mathscr{F}$-可测的，且($\mathrm{mod}P$)有限. 注意，

$$\int_A r_\theta \mathrm{d}P = \int_A r_\theta \cdot f \cdot h \mathrm{d}\mu = \int_{A\cdot\{f>0\}} g_\theta \cdot h \mathrm{d}\mu$$

$$= P_\theta(A) \ (\forall\ A \in \mathscr{F})$$

$$\Rightarrow r_\theta = \frac{\mathrm{d}P_\theta}{\mathrm{d}P} \ (\mathrm{mod}P).$$

这表明：$\frac{\mathrm{d}P_\theta}{\mathrm{d}P}$ 是 $\mathscr{G}$-可测的，且($\mathrm{mod}P$)有限. 按引理 3.5.1，当以 P 代 μ，1 代 h 时，便可得结论：$\mathscr{G}$ 关于(P_θ)是充足的.

必要性 假设 $\mathscr{G}$ 关于(P_θ)是充足子 σ-代数. 按定义 3.5.1，$\exists$ 二元函数 $p(A,\omega)$ $(A\times\omega\in\mathscr{F}\times\Omega)$，使得

$$p(A,\omega) = P_\theta(A/\mathscr{G}) \ (\mathrm{mod}P_\theta) \ (\forall\ \theta \in \Theta).$$

按上述证明中定义 P，则对 $\forall\ G\in\mathscr{G}$，有

$$\int_G p(A,\omega)P(\mathrm{d}\omega) = \sum_{n=1}^{\infty} 2^{-n} \int_G p(A,\omega) P_{\theta_n}(\mathrm{d}\omega)$$

$$= \sum_{n=1}^{\infty} 2^{-n} \int_G P_{\theta_n}(A/\mathscr{G}) \mathrm{d}P_{\theta_n}$$

$$= \sum_{n=1}^{\infty} 2^{-n} P_{\theta_n}(AG) = P(AG),$$

即 $p(A,\omega) = P(A/\mathscr{G})$ ($\mathrm{mod}P$).

现将 P 并入 $(P_\theta)_{\theta\in\Theta}$ 中，则 $\mathscr{G}$ 关于此扩大族仍是充足的. 按引理 3.5.2 便可得所要求分解.

设 Θ 为一个实数子集. r.v. Z 为 θ 的一个估计. 若 $E_\theta Z=\theta$ $(\forall\ \theta\in\Theta)$，则说 Z 为一个**无偏估计**. 方差值 $E_\theta(Z-\theta)^2$ 为估计 Z 所引起的**偏差度量**.

引理 3.5.4 若 $\mathscr{G}$ 关于 $(P_\theta)_{\theta\in\Theta}$ 是充足子 σ-代数. 则 $\exists\ \mathscr{F}$ 上的概率 P，$\dashv$ 对任意 P_θ-可积的 r.v. X，有

$$E(X/\mathscr{G}) = E_\theta(X/\mathscr{G})\ (\mathrm{mod} P_\theta)\ (\forall\ \theta\in\Theta).$$

证 从上面的引理之证中即可得此结论.

定理 3.5.2（Rao-Blackwell 不等式） 设 $\mathscr{G}$ 关于 (P_θ) 是充足的子 σ-代数，且 (P_θ) 受 σ-有限测度 μ 的控制. Z 为 θ 的一个估计，且 $E_\theta(Z-\theta)^2<\infty$ $(\forall\ \theta\in\Theta)$. 则 $\exists\ \mathscr{F}$ 上的概率测度 P，使得

$$E_\theta(E(Z/\mathscr{G}) - \theta)^2 \leqslant E_\theta(Z-\theta)^2 (\forall\ \theta\in\Theta).$$

若 Z 无偏，即 $E_\theta Z=\theta$ $(\forall\ \theta\in\Theta)$，则 $E(Z/\mathscr{G})$ 亦是无偏的.

证 按引理 3.5.4，$\exists\ \mathscr{F}$ 上的概率测度 P，使得

$$\begin{aligned}
&E(Z/\mathscr{G}) = E_\theta(Z/\mathscr{G})\ (\mathrm{mod} P_\theta)\ (\forall\ \theta\in\Theta)\\
\Rightarrow\ &E_\theta(E(Z/\mathscr{G}) - \theta)^2 = E_\theta(E_\theta(Z/\mathscr{G}) - \theta)^2\\
&\qquad = E_\theta(E_\theta(Z/\mathscr{G})^2 - 2\theta\cdot E_\theta(Z/\mathscr{G}) + \theta^2)\\
&\qquad \leqslant E_\theta Z^2 - 2\theta E_\theta Z + \theta^2 = E_\theta(Z-\theta)^2.
\end{aligned}$$

这是定理的第一个结论. 第二个结论显然.

例 3.5.2 继续讨论例 3.5.1. 在此例中已证明 $\sigma(T)$ 关于 (P_θ) 是充足的. 现在，先考查估计 $2\overline{X}=2\left(\frac{1}{k}\sum_{i=1}^{k}X_i\right)$. 其均值为 $E_\theta 2\overline{X}=\int_0^\theta 2u\cdot\mathrm{d}\left(\frac{u}{\theta}\right)=\theta$. 因此，$2\overline{X}$ 为 θ 的无偏估计. 考查另一个估计 $T^*=\frac{k+1}{k}T$，则 $T^*=E_\theta(2\overline{X}/T)\ (\mathrm{mod} P_\theta)$（见例 3.5.1）. 于是，$T^*$ 亦为 θ 的一个无偏估计，且

$$E_\theta(T^*-\theta)^2 < E_\theta(2\overline{X}-\theta)^2 (\forall\ \theta\in\Theta)，其中\ k>1.$$

现在，我们要证明：偏差度量 $E_\theta(T^*-\theta)^2$ 在 θ 的全体 P_θ-平方可积

无偏估计类中是最小的. 为此,设 Z 为 θ 的任意 P_θ-平方可积无偏估计,则按定理 3.5.2,有

$$E_\theta(E(Z/T)-\theta)^2 \leqslant E_\theta(Z-\theta)^2 \quad (\forall\ \theta \in \Theta).$$

注意,$E_\theta(E(Z/T)-\theta)=0$. 于是,

$$\begin{aligned} E_\theta(T^*-\theta)^2 &= E_\theta(T^*-E(Z/T))^2 \\ &\quad + 2E_\theta[(T^*-E(Z/T))\cdot(E(Z/T)-\theta)] \\ &\quad + E_\theta(E(Z/T)-\theta)^2 \\ &\leqslant E_\theta(T^*-E(Z/T))^2 + E_\theta(Z-\theta)^2. \end{aligned}$$

若能证明: $T^*-E(Z/T)=0\ (\mathrm{mod}P_\theta)\ (\forall\ \theta>0)$,则

$$E_\theta(T^*-\theta)^2 \leqslant E_\theta(Z-\theta)^2 \quad (\forall\ \theta \in \Theta).$$

此不等式正是我们所要的结论.

令 $f(T)=T^*-E(Z/T)$,则有 $E_\theta f(T)=0\ (\forall\ \theta\in\Theta)$. 另一方面, $E_\theta f(T)=\int_0^\theta f(t)\mathrm{d}\left(\frac{t^k}{\theta^k}\right)=\frac{k}{\theta^k}\int_0^\theta f(t)\cdot t^{k-1}\mathrm{d}t\ (\forall\ \theta>0)$. 故

$$\int_0^\theta f(t)\cdot t^{k-1}\mathrm{d}t=0\ (\forall\ \theta>0).$$

由此即得 $f(t)=0\ (\mathrm{mod}\lambda_1)$. 从而, $f(T)=0\ (\mathrm{mod}P_\theta)\ (\forall\ \theta\in\Theta)$. 故 $E_\theta(T^*-E(Z/T))^2=0\ (\forall\ \theta\in\Theta)$.

第四章　离散参数鞅

鞅论是在条件期望理论基础上发展起来的一个概率论分支.本章只介绍离散参数鞅.鞅论这一分支不论在理论上还是在应用上都有广泛的发展前景.普遍受到人们的重视.本章的内容是进入鞅论的重要一步.我们将首先介绍取离散值的停时,然后介绍鞅概念、鞅性质及某些典型的应用.

§4.1　停　　时

设 $\mathrm{N}=\{0,1,2,\cdots\}$,$\overline{\mathrm{N}}=\mathrm{N}\cup\{+\infty\}$.若 $\mathscr{F}$ 中的子 σ-代数族 $(\mathscr{F}_n,n\in\mathrm{N})$具有性质:$\mathscr{F}_n\subseteq\mathscr{F}_{n+1}(\forall\ n\in\mathrm{N})$,则称$(\mathscr{F}.)$为 $\mathscr{F}$ 中的 σ-代数流.$\overline{\mathscr{N}}$ 表示 $\overline{\mathscr{N}}$ 中子集之全体.$(\Omega,\mathscr{F},\mathrm{P})$为某个概率空间.$\xi=(\xi_n,n\in\mathrm{N})$是定义在此空间上的随机变量列.

定义 4.1.1　若对每个 $n\in N$,ξ_n 是 $\mathscr{F}_n$-可测的,则称 ξ 适应于 σ-代数流$(\mathscr{F}.)$.

按此定义,若记 $\mathscr{F}_n^\xi=\sigma(\xi_0,\xi_1,\cdots,\xi_n)$,则$(\mathscr{F}_\cdot^\xi)$为 $\mathscr{F}$ 中的一个 σ-代数流.显然,ξ 是适应于$(\mathscr{F}_\cdot^\xi)$的.这种情形常称为**自适应**.

定义 4.1.2　设 τ 是$(\Omega,\mathscr{F})\to(\overline{\mathrm{N}},\overline{\mathscr{N}})$的可测映射.若对$\forall\ n\in\mathrm{N}$,有

$$(\tau=n)\triangleq\{\omega:\tau(\omega)=n\}\in\mathscr{F}_n,$$

则称 τ 为关于$(\mathscr{F}.)$的停时.若 $P(\tau<\infty)=1$,则称 τ 为有限停时.若$\exists\ n_0\in\mathrm{N}$,$\ni P(\tau\leqslant n_0)=1$,则称 τ 为有界停时.

注 1　从定义中可以看出:停时 τ 涉及到三个要素:$\overline{\mathrm{N}}$,$(\mathscr{F}.)$,P.当$(\mathscr{F}.)$改变时,τ 可以不是停时.当 P 改变时,有限停时可以变成非有限停时.

定义 4.1.3　设 τ 是$(\Omega,\mathscr{F})\to(\overline{\mathrm{N}},\overline{\mathscr{N}})$的可测映射.若

$$(\tau = n) \in \mathscr{F}_{n-1}(\forall\ n \in \mathrm{N}),$$

其中 $\mathscr{F}_{-1}=\mathscr{F}_0$，则称 τ 为可料停时.

引理 4.1.1 设 $(\tau_k, k\geqslant 1)$ 为(可料)停时列，则 $\tau_1+\tau_2$，$\inf\limits_k \tau_k$，$\sup\limits_k \tau_k$，$\underline{\lim}\tau_k$ 和 $\overline{\lim}\tau_k$ 均为(可料)停时.

证 $(\tau_1+\tau_2=n)=\bigcup\limits_{k=0}^{n}(\tau_1=k)(\tau_2=n-k)\in\mathscr{F}_n(\mathscr{F}_{n-1})$.

$$(\inf_k \tau_k\geqslant n+1)=\bigcap_{k=1}^{\infty}(\tau_k\geqslant n+1)\in\mathscr{F}_n(\mathscr{F}_{n-1}).$$

$$(\sup_k \tau_k\leqslant n)=\bigcap_{k=1}^{\infty}(\tau_k\leqslant n)\in\mathscr{F}_n(\mathscr{F}_{n-1}).$$

$$(\underline{\lim}\tau_k\leqslant n)=(\sup_m \inf_{k\geqslant m}\tau_k\leqslant n)\in\mathscr{F}_n(\mathscr{F}_{n-1}).$$

$$(\overline{\lim}\tau_k\geqslant n+1)=(\inf_m \sup_{k\geqslant m}\tau_k\geqslant n+1)\in\mathscr{F}_n(\mathscr{F}_{n-1}).$$

易证："$(\tau=n)\in\mathscr{F}_n(\mathscr{F}_{n-1})(\forall\ n\in N)$"等价于"$(\tau\leqslant n)\in\mathscr{F}_n(\mathscr{F}_{n-1})(\forall\ n\in N)$". 故按定义 4.1.2，即得所要的结论.

引理 4.1.2 设 r. v. 列 $\xi=(\xi_n, n\in\mathrm{N})$ 适应于 $(\mathscr{F}.)$，τ 为停时. 令

$$\xi_\tau=\begin{cases}\xi_n, & \text{当 } \omega\in(\tau=n) \text{ 时}(\forall\ n\in\mathrm{N}),\\ 0, & \text{当 } \omega\in(\tau=\infty) \text{ 时}.\end{cases}$$

则 ξ_τ 为 r. v.

证 设 $B\in\mathscr{B}$，则

$$\{\xi_\tau\in B\}=\bigcup_{n=0}^{\infty}\{\xi_n\in B\}\cdot(\tau=n)\cup\{\xi_\infty\in B\}\cdot(\tau=\infty).$$

此式中的每一项都是 $\mathscr{F}$ 中的元素，故 $\{\xi_\tau\in B\}\in\mathscr{F}$.

引理 4.1.3 设 $\xi=(\xi_n, n\in\mathrm{N})$ 为适应的 r. v. 列. 定义：

$$\tau_B=\inf\{n\in\mathrm{N}:\xi_n\in B\},\inf\{\varnothing\}=\infty$$

其中 $B\in\mathscr{B}$，则 τ_B 为停时.

证 对 $\forall\ n\geqslant 1$，有

$$(\tau_B=n)=\{\omega:\xi_k\notin B, 0\leqslant k\leqslant n-1;\xi_n\in B\}\in\mathscr{F}_n(n\geqslant 1),$$

$$(\tau_B=0)=\{\omega:\xi_0\in B\}\in\mathscr{F}_0.$$

故 τ_B 为停时.

注 2 令 $Y=\sup\limits_k \xi_k$,则对 $B=[b,\infty)(b\in R)$,有 $\{Y\in B\}=\bigcup\limits_{n=0}^{\infty}(\tau_B=n)=(0\leqslant\tau_B<\infty)$. 停时与适应序列 ξ 之间的这个关系是很有用的.

定义 4.1.4 设 τ 为停时,定义:

$$\mathscr{F}_\tau=\{A\in\mathscr{F}:A\cap(\tau=n)\in\mathscr{F}_n(\forall\, n\in\mathrm{N})\},$$

则称 $\mathscr{F}_\tau$ 是与 τ 相联的 σ-代数. 一种等价的定义是将 $\mathscr{F}$ 由 $\mathscr{F}_\infty=\sigma(\bigcup\limits_{n=0}^{\infty}\mathscr{F}_n)$代替.

引理 4.1.4 设 τ、τ_1、τ_2 均为停时.

(i) τ 是 $\mathscr{F}_\tau$-可测的,即 $\sigma(\tau)\subseteq\mathscr{F}_\tau$;

(ii) $\tau_1\leqslant\tau_2\Rightarrow\mathscr{F}_{\tau_1}\subseteq\mathscr{F}_{\tau_2}$;

(iii) $\mathscr{F}_\tau\cap(\tau=n)=\mathscr{F}_n\cap(\tau=n)(\forall\, n\in\mathrm{N})$;

(iv) $A=(\tau_1<\tau_2)$或$(\tau_1>\tau_2)$或$(\tau_1=\tau_2)\Rightarrow A\in\mathscr{F}_{\tau_1\wedge\tau_2}=\mathscr{F}_{\tau_1}\cap\mathscr{F}_{\tau_2}$.

证 结论(i)显然. 如果 $\tau_1\leqslant\tau_2$,则对$\forall\, A\in\mathscr{F}_{\tau_1}$,有

$$A\cap(\tau_1\leqslant n)\in\mathscr{F}_n(\forall\, n\in\mathrm{N}).$$

$$\Rightarrow A\cap(\tau_2=n)=A\cap(\tau_1\leqslant n)\cap(\tau_2=n)\in\mathscr{F}_n(\forall\, n\in\mathrm{N}).$$

故 $\mathscr{F}_{\tau_1}\subseteq\mathscr{F}_{\tau_2}$,此即结论(ii). 设 $A\in\mathscr{F}_\tau$,则 $A\cdot(\tau=n)\in\mathscr{F}_n$. 从而,$A\cdot(\tau=n)\cdot(\tau=n)\in\mathscr{F}_n\cap(\tau=n)$. $\Rightarrow\mathscr{F}_\tau\cap(\tau=n)\subset\mathscr{F}_n\cap(\tau=n)$. 现在,让 $A\in\mathscr{F}_n\cap(\tau=n)$,则 $A\cap(\tau=k)=\begin{cases}A,k=n,\\ \varnothing,k\neq n.\end{cases}\ \forall\, k\in\mathrm{N}$. 从而,$A\cap(\tau=k)\in\mathscr{F}_k(\forall\, k\in\mathrm{N})$. $\Rightarrow A\in\mathscr{F}_\tau\Rightarrow\mathscr{F}_n\cap(\tau=n)\subset\mathscr{F}_\tau\cap(\tau=n)$. 综合起来得结论(iii). 为证(iv),考虑到对称性,只需证明:$A=(\tau_1<\tau_2)\in\mathscr{F}_{\tau_1\wedge\tau_2}$. 为此,对$\forall\, n\in\mathrm{N}$,不难得:

$$(\tau_1<\tau_2)(\tau_2=n)=(\tau_1<n)(\tau_2=n)\in\mathscr{F}_n,$$

$$(\tau_1<\tau_2)(\tau_1=n)=(\tau_1=n)(\tau_2>n)\in\mathscr{F}_n$$

$$\Rightarrow(\tau_1<\tau_2)\in\mathscr{F}_{\tau_1}\cap\mathscr{F}_{\tau_2}.$$

假定 $A\in\mathscr{F}_{\tau_1}\cap\mathscr{F}_{\tau_2}$，则

$$A(\tau_1\wedge\tau_2=n)=A(\tau_1=n)(\tau_2\geqslant n)\cup A(\tau_1\geqslant n)(\tau_2=n)\in\mathscr{F}_n.$$

$\Rightarrow A\in\mathscr{F}_{\tau_1\wedge\tau_2}$，即 $\mathscr{F}_{\tau_1}\cap\mathscr{F}_{\tau_2}\subseteq\mathscr{F}_{\tau_1\wedge\tau_2}$.

反之，假定 $A\in\mathscr{F}_{\tau_1\wedge\tau_2}$，则

$$\begin{aligned}A(\tau_1=n)&=A(\tau_1=n)(\tau_2\geqslant n)\cup A(\tau_1=n)(\tau_2\leqslant n)\\&=A(\tau_1\wedge\tau_2=n)(\tau_1=n)\cup A(\tau_1\wedge\tau_2\leqslant n)\\&\quad(\tau_1=n)\in\mathscr{F}_n\end{aligned}$$

$\Rightarrow A\in\mathscr{F}_{\tau_1}$. 同理，$A\in\mathscr{F}_{\tau_2}$.

$$\Rightarrow\mathscr{F}_{\tau_1\wedge\tau_2}\subseteq\mathscr{F}_{\tau_1}\cap\mathscr{F}_{\tau_2}.$$

综上分析，便得结论(iv).

注 3 为什么不考虑 $\sigma(\tau)$，而把注意力放在 $\mathscr{F}_\tau$ 上. 这个问题可以从下面的引理得到答案.

引理 4.1.5 设 ξ 为适应于 $(\mathscr{F}.)$ 的 r.v 列. τ 为停时，ξ_τ 按引理 4.1.2 定义，则 ξ_τ 是 $\mathscr{F}_\tau$-可测的.

证 按引理 4.1.2，$A=\{\xi_\tau\in B\}\in\mathscr{F}(\forall\ B\in\mathscr{B})$. 又

$$A\cdot(\tau=n)=\{\xi_n\in B\}\cdot(\tau=n)\in\mathscr{F}_n(\forall\ n\in\mathrm{N}).$$

故 $A\in\mathscr{F}_\tau$. 此即所要的结论.

定义 4.1.5 设 $\xi=(\xi_n,n\in\mathrm{N})$ 为 r.v. 列. 若 ξ_n 是 $\mathscr{F}_{n-1}$-可测的 $(\forall\ n\in\mathrm{N})(\mathscr{F}_{-1}=\mathscr{F}_0)$，则称 ξ 为可料序列.

任意可料序列必为适应序列，但反之不然.

引理 4.1.6 若 ξ 为可料序列，τ 为可料停时，则 ξ_τ 是 $\mathscr{F}_\tau^*$-可测的，其中

$$\mathscr{F}_\tau^*=\{A\in\mathscr{F}:A\cdot(\tau=n)\in\mathscr{F}_{n-1}(\forall\ n\in\mathrm{N})\}.$$

证 类似于引理 4.1.5 之证.

引理 4.1.7 若 ξ 为可料序列，$(\tau_k,k\geqslant1)$ 为上升的可料停时列. 令 $Y_k=\xi_{\tau_k}$，$\mathscr{F}_k^*=\mathscr{F}_{\tau_k}^*(\forall\ k\geqslant1)$，则 $Y=(Y_k,k\geqslant1)$ 为适应于 $(\mathscr{F}_k^*,k\geqslant1)$ 的 r.v. 列.

证 由引理 4.1.6 立即可推出此结论.

§4.2 鞅的定义和例

设$(\Omega,\mathscr{F},P)$为概率空间. $(\mathscr{F}_n,n\in N)$为$\mathscr{F}$中σ-代数流.

定义 4.2.1 设$\xi=(\xi_n,n\in \mathrm{N})$为$(\Omega,\mathscr{F},P)$上的广义 r. v. 列,且适应于$(\mathscr{F}.)$.

(i) 若 $\xi_n\in L'\triangleq L'(\Omega,\mathscr{F},P)$,且 $E(\xi_{n+1}/\mathscr{F}_n)=(\geqslant,\leqslant)\xi_n(\mathrm{mod}P)(\forall\ n\in \mathrm{N})$,则称$\xi$为关于$(P,\mathscr{F}.)$的鞅(下鞅,上鞅).

(ii) 若$E(\xi_{n+1}^+/\mathscr{F}_n)\wedge E(\xi_{n+1}^-/\mathscr{F}_n)<\infty(\mathrm{mod}P)(\forall\ n\in \mathrm{N})$,且

$$E(\xi_{n+1}/\mathscr{F}_n)=(\geqslant,\leqslant)\xi_n(\mathrm{mod}P)\ (\forall\ n\in \mathrm{N}),$$

则称ξ为关于$(P,\mathscr{F}.)$的广义鞅(广义下鞅,广义上鞅).

这里说到"广义 r. v."其含意是可以取$\pm\infty$. 而"r. v. "可以理解为取$\pm\infty$的概率为 0. 当$\xi_{n+1}\in L'$时,它关于$\mathscr{F}_n$的条件期望为 r. v. 然而,对任意 r. v. ξ_{n+1},它关于$\mathscr{F}_n$的条件期望不必为 r. v. ,即为广义 r. v. 随机元$\xi=(\xi_n,n\in \mathrm{N})$的鞅性质与概率$P$和$\sigma$-代数流$(\mathscr{F}.)$有关,$P$和$(\mathscr{F}.)$的改变均可导致鞅性质的改变. 以符号$M(P,\mathscr{F}.)$,$M_s(P,\mathscr{F}.)$,$M^s(P,\mathscr{F}.)$分别表示关于$(P,\mathscr{F}.)$的鞅、下鞅、上鞅序列之全体. $A_s(\mathscr{F}.)$表示适应于$(\mathscr{F}.)$的 r. v. 列之全体.

引理 4.2.1 设$\xi\in M(P,\mathscr{F}.)(M_s(P,\mathscr{F}.))$,$\tau$为停时.

令 $\mathscr{F}_n^\tau=\mathscr{F}_{\tau\wedge n}$,$\xi_n^\tau=\xi_{\tau\wedge n}(\forall\ n\in \mathrm{N})$,则

$$\xi^\tau=(\xi_n^\tau,n\in \mathrm{N})\in M(P,\mathscr{F}^\tau.)(M_s(P,\mathscr{F}^\tau.)).$$

证 按引理 4.1.1,对$\forall\ n\in N$,$\tau\wedge n$为停时. 按引理 4.1.5,$\xi_{\tau\wedge n}$是$\mathscr{F}_{\tau\wedge n}$-可测的. 按条件$\xi\in M(P,\mathscr{F}.)$,有

$$E(\xi_{\tau\wedge(n+1)}/\mathscr{F}_{n\wedge\tau})1_{(\tau>n)}=E(1_{(\tau>n)}\xi_{n+1}/\mathscr{F}_n)$$
$$=E(\xi_{n+1}/\mathscr{F}_n)1_{(\tau>n)}=(\geqslant)\xi_n^\tau\cdot 1_{(\tau>n)}(\mathrm{mod}P).$$

$$E(\xi_{\tau\wedge(n+1)}/\mathscr{F}_{n\wedge\tau})1_{(\tau\leqslant n)}=\sum_{k=0}^{n}E(\xi_{\tau\wedge(n+1)}/\mathscr{F}_{(\tau\wedge n)})1_{(\tau=k)}$$

$$= \sum_{k=0}^{n} E(\xi_k/\mathscr{F}_k)1_{(\tau=k)} = \xi_n^\tau \cdot 1_{(\tau\leqslant n)}.$$

$$\Rightarrow E(\xi_{\tau\wedge(n+1)}/\mathscr{F}_{\tau\wedge n}) = (\geqslant)\xi_{\tau\wedge n}(\mathrm{mod}P).$$

故 $\xi^\tau \in M(P,\mathscr{F}^\tau.)(M_s(P,\mathscr{F}^\tau.))$.

注 1 这个结论是 Doob 停时代换不变性的一种特别形式. 它可以稍微变得更一般些：设$(\tau_k,k\in\mathrm{N})$为上升的停时列，且每个 τ_k 有界，则引理 4.2.1 仍成立.

定义 4.2.2 设$\xi\in A_s(\mathscr{F}.)$. 若$\exists$ 停时列$(\tau_k,k\geqslant1)$满足下列要求：

(i) $\tau_k\uparrow\infty$ (a.s)；

(ii) $\xi^{\tau_k}=(\xi_{n\wedge\tau_k}1_{(\tau_k>0)},\mathscr{F}_{n\wedge\tau_k},n\in\mathrm{N})\in M(P,\mathscr{F}^{\tau_k}.)(M_s(P,\mathscr{F}^{\tau_k}.))$,

则称 ξ 为关于$(\mathscr{F}.)$的局部鞅(局部下鞅).

定义 4.2.3 设$\xi\in A_s(\mathscr{F}.)$，$\nu=(\nu_n,n\in\mathrm{N})$为可料序列.

令
$$\nu\cdot\xi=((\nu\cdot\xi)_n,\mathscr{F}_n,n\in\mathrm{N}),$$
其中
$$\Delta\xi_n=\xi_n-\xi_{n-1},\xi_{-1}=0,$$
$$(\nu\cdot\xi)_n=(\nu\cdot\xi)_{n-1}+\nu_n\cdot\Delta\xi_n(\forall\ n\in\mathrm{N}).$$

则称 $\nu\cdot\xi$ 为 ξ 的 ν-变换.

定理 4.2.1 设$\xi\in A_s(\mathscr{F}.)$. 下列条件等价：

(a) ξ 为局部鞅；

(b) ξ 为广义鞅；

(c) ξ 为某鞅的 ν-变换.

证 设 ξ 为局部鞅. 按定义 4.2.2，存在停时列$(\tau_k,k\geqslant1)$，$\tau_k\uparrow\infty$，$\ni\xi^{\tau_k}=(1_{(\tau_k>0)}\cdot\xi_{n\wedge\tau_k})_{n\in N}\in M(P,\mathscr{F}^{\tau_k}.)$ $(\forall\ k\geqslant1)$

$$\Rightarrow E(1_{(\tau_k>0)}|\xi_{n\wedge\tau_k}|)<\infty(\forall\ n\in\mathrm{N},k\geqslant1)$$

$$\Rightarrow E(|\xi_{n+1}|/\mathscr{F}_n)1_{(\tau_k>n)} = E(|\xi_{n+1}|1_{(\tau_n>n)}/\mathscr{F}_n)$$
$$= E(|\xi_{(n+1)\wedge\tau_k}|1_{(\tau_k>n)}/\mathscr{F}_n)$$
$$\leqslant E(|\xi_{(n+1)\wedge\tau_k}|\cdot1_{(\tau_k>0)}/\mathscr{F}_n)<\infty\ (\mathrm{mod}P).$$

注意，$(\tau_k>n)\uparrow\Omega$ 当 $k\uparrow\infty$ 时. 于是，对$\forall\ n\in\mathrm{N}$，有

$$E(|\xi_{n+1}|/\mathscr{F}_n)<\infty\ (\mathrm{mod}P).$$

从而，$E(\xi_{n+1}/\mathscr{F}_n)$为 r. v.. 于是

$$\begin{aligned}\xi_n\cdot 1_{(\tau_k>n)}&=1_{(\tau_k>n)}\xi_{n\wedge\tau_k}\cdot 1_{(\tau_k>0)}\\&=1_{(\tau_k>n)}E(\xi_{(n+1)\wedge\tau_k}1_{(\tau_k>0)}/\mathscr{F}_{n\wedge\tau_k})\\&=E(\xi_{(n+1)\wedge\tau_k}\cdot 1_{(\tau_k>n)}/\mathscr{F}_n)\\&=1_{(\tau_k>n)}E(\xi_{n+1}/\mathscr{F}_n)\ (\mathrm{a.s}).\end{aligned}$$

在此式中，让 $k\uparrow\infty$，则可得：

$$\xi_n=E(\xi_{n+1}/\mathscr{F}_n)\ (\mathrm{a.s})\ (\forall\ n\in\mathrm{N}).$$

故 ξ 为广义鞅，即$(a)\Rightarrow(b)$.

现在设 ξ 为广义鞅. 这时，$E(|\Delta\xi_{n+1}|/\mathscr{F}_n)$为 r. v. 事实上，按广义鞅之定义，不妨设 $E(\xi_{n+1}^-/\mathscr{F}_n)$为 r. v. 按条件 $\xi\in A_s(\mathscr{F}.)$，

$$E(\xi_{n+1}^+/\mathscr{F}_n)=\xi_n+E(\xi_{n+1}^-/\mathscr{F}_n)$$

为 r. v. 于是，

$$E(|\Delta\xi_{n+1}|/\mathscr{F}_n)\leqslant E(|\xi_{n+1}|/\mathscr{F}_n)+|\xi_n|<\infty\ (\mathrm{a.s}).$$

令

$$\nu_n=\begin{cases}E(|\Delta\xi_n|/\mathscr{F}_{n-1})\neq 0,\\ 1,\text{当 }E(|\Delta\xi_n|/\mathscr{F}_{n-1})=0\text{ 时}.\end{cases}\quad(\forall\ n\in\mathrm{N})$$

$$\eta_n=\sum_{k=0}^{n}\frac{1}{\nu_k}\Delta\xi_k(\forall\ n\in\mathrm{N}).\ (\text{规定 }\xi_{-1}=0,\mathscr{F}_{-1}=\mathscr{F}_0).$$

则 $\xi_n=(\nu\cdot\eta)_n(\forall\ n\in\mathrm{N})$. 这是因为

$$(\nu\cdot\eta)_n=\nu_0\eta_0+\sum_{k=1}^{n}\nu_k\cdot\Delta\eta_k=\xi_n(\forall\ n\in\mathrm{N}).$$

显然，ν 为可料序列. 而

$$\begin{aligned}E|\eta_n|&\leqslant\sum_{k=0}^{n}E\left(\frac{|\Delta\xi_k|}{\nu_k}\right)\\&=\sum_{k=0}^{n}E\left(\frac{1}{\nu_k}E(|\Delta\xi_k|/\mathscr{F}_{k-1})\right)\leqslant n+1<\infty.\end{aligned}$$

即 $\eta_n\in L^1$. 又

$$E(\Delta\eta_n/\mathscr{F}_{n-1})=E\left(\frac{1}{\nu_k}\Delta\xi_n/\mathscr{F}_{n-1}\right)=\frac{1}{\nu_n}E(\Delta\xi_n/\mathscr{F}_{n-1})=0.$$

故 η 为鞅. 从而,得$(b)\Rightarrow(c)$.

最后,设 ξ 为某鞅 η 的 ν-变换,即 $\xi_n=(\nu\cdot\eta)_n$. 令

$$\tau_k=\inf\{n\in N:|\nu_0|\leqslant k,|\nu_{n+1}|>k\},\inf\{\varnothing\}=\infty.$$

则由 ν 的可料性知:τ_k 为停时,且 $\tau_k\uparrow\infty$当 $k\uparrow\infty$时. 由 $\eta\in M(P,\mathscr{F}.)$及 τ_k 之定义可知:$\xi_{n\wedge\tau_k}\in L'$. 记

$$\Delta\xi_n^{\tau_k}=\xi_{n\wedge\tau_k}-\xi_{(n-1)\wedge\tau_k}.$$

则$E(1_{(\tau_k>0)}\Delta\xi_{n+1}^{\tau_k}/\mathscr{F}_{n\wedge\tau_k})=E(1_{(\tau_k>0)}\nu_{(n+1)\wedge\tau_k}\Delta\eta_{n+1}^{\tau_k}/\mathscr{F}_{n\wedge\tau_k})$

$$=1_{(\tau_k>0)}\nu_{(n+1)\wedge\tau_k}E(\Delta\eta_{n+1}^{\tau_k}/\mathscr{F}_{n\wedge\tau_k})=0.$$

故 $\xi^{\tau_k}1_{(\tau_k>0)}\in M(P,\mathscr{F}.^{\tau_k})$ $(\forall\ k\geqslant1)$. 由此即得$(c)\Rightarrow(a)$.

例 4.2.1 似然比.

设$(\Omega,\mathscr{F},P)$为概率空间. $Y=(Y_n,n\geqslant1)$为$(\Omega,\mathscr{F})$上的随机序列. $\mathscr{F}_n=\sigma(Y_1,Y_2,\cdots,Y_n)$. Q 为$(\Omega,\mathscr{F})$上的另一个概率测度,且 $Q_n\ll P_n$,其中 $Q_n=Q|\mathscr{F}_n$,$P_n=P|\mathscr{F}_n$. $(\forall\ n\geqslant1)$. 令 $X_n=\frac{\mathrm{d}Q_n}{\mathrm{d}P_n}$,则 $X=(X_n,n\geqslant1)$具有下列性质:

(1) $X\in M(P,\mathscr{F}.)$. 事实上,对$\forall\ A\in\mathscr{F}_n$,有

$$\int_A X_{n+1}\mathrm{d}P=\int_A\frac{\mathrm{d}Q_{n+1}}{\mathrm{d}P_{n+1}}\mathrm{d}P_{n+1}=Q_{n+1}(A)=Q_n(A)$$

$$=\int_A X_n\mathrm{d}P_n=\int_A X_n\mathrm{d}P.$$

注 2 一般地说,$X\notin M(Q,\mathscr{F}.)$

(2) 设 Q_n^*,P_n^* 由$(Y_1,Y_2,\cdots,Y_n)$分别相应于 Q_n,P_n 所诱导的$(R^n,\mathscr{B}^n)$上的概率分布,且 $Q_n^*\ll\lambda_n$,$P_n^*\ll\lambda_n$,这里 λ_n 为 n-维 Lebesque 测度. 则

$$X_n=\frac{\mathrm{d}Q_n}{\mathrm{d}P_n}=\frac{\mathrm{d}Q_n^*}{\mathrm{d}P_n^*}(Y_1,\cdots,Y_n)=\frac{q_n(Y_1,\cdots,Y_n)}{p_n(Y_1,\cdots,Y_n)},$$

其中
$$q_n=\frac{\mathrm{d}Q_n^*}{\mathrm{d}\lambda_n},\quad p_n=\frac{\mathrm{d}P_n^*}{\mathrm{d}\lambda_n}.$$

(3) 假定 Y 关于 P、Q 均是独立、同分布的. 则

$$X_n = \prod_{i=1}^{n} \frac{q(Y_i)}{p(Y_i)},$$

其中 $p(y)$、$q(y)$ 为 Y_1 分别相应于 P、Q 的分布密度.

(4) 在(3)中的假设下，若 p、q 不恒等$(\bmod \lambda_1)$，则 $P\perp Q$，即 X 不收敛. 事实上，p、q 不恒等$(\bmod \lambda_1)$表明：$\exists\ H\in\mathscr{B}$，$\ni P(Y_n\in H)\neq Q(Y_n\in H)$. 令 $Z_n=1_{\{Y_n\in H\}}$. 则按强大数定律，有

$$\frac{1}{n}\sum_{k=1}^{n} Z_k \to P(Y_1 \in H)\ (\bmod P);$$

$$\frac{1}{n}\sum_{k=1}^{n} Z_k \to Q(Y_1 \in H)\ (\bmod Q).$$

令 $A=\left\{\omega: \frac{1}{n}\sum_{k=1}^{n} Z_k \to P(Y_1\in H)\right\}$，则由 $P(Y_1\in H)\neq Q(Y_1\in H)$知：$P(A)=1$，$Q(A^c)=1$，即 $P\perp Q$.

注 3 最后这一点表明，即使对$\forall\ n\geqslant 1$，有 $Q_n\ll P_n$. 仍不能保证：$Q\ll P$.

例 4.2.2 分枝过程.

设$(N_{nk}: n\geqslant 1, k\geqslant 1)$为 r. v. 族，$N_{nk}$取值于 N，且对$\forall\ n\geqslant 1$，$(N_{nk}, k\geqslant 1)$是独立同分布的. 定义 $Z_0=1$，

$$Z_n(\omega) = \begin{cases} N_{n1}(\omega) + \cdots + N_{nZ_{n-1}(\omega)}(\omega), & \text{当 } Z_{n-1}(\omega) \neq 0 \text{ 时}, \\ 0, & \text{否则}. \end{cases}$$

则 $Z=(Z_n, n\in\mathrm{N})$为分枝过程. r. v. N_{nk}表示第 n 代中第 k 个子群体的数目，Z_n 表示第 n 代各子群体数目的总和.

设 $EN_{nk}=m\neq 0$，则

$$E(Z_n/Z_{n-1}) = E\left(\sum_{k=1}^{Z_{n-1}} N_{nk}/Z_{n-1}\right) = m \cdot Z_{n-1}.$$

令 $X_n=m^{-n}\cdot Z_n(n\in N)$，则 $X=(X_n, n\in N)$为自生鞅.

事实上，令

$$\mathscr{F}_n^X = \sigma(X_0, X_1, \cdots, X_n) = \sigma(Z_0, Z_1, \cdots, Z_n),$$

则 $$E(X_{n+1}/\mathscr{F}_n^X) = E(m^{-(n+1)}Z_{n+1}/Z_n)$$
$$= m^{-(n+1)}Z_n \cdot m = m^{-n} \cdot Z_n = X_n.$$
注意,分枝过程具有性质:若已知当前一代,则未来的发展独立于过去的历史状态.即该过程具有 Markov 性.

例 4.2.3 博弈问题.

甲、乙双方对弈,η_n 表示甲方在第 n 轮对弈中胜负的指示量.$\eta_n=1$ 和 $\eta_n=-1$ 分别表示甲方胜和负.ξ_n 表示在 n 轮对弈后的总结果.ν_n 表示甲方在第 n 轮对弈中所下的赌注.ν_0 表示甲方所拥有的初始赌资.$\eta_0\equiv1$.则

$$\xi_n = \sum_{k=0}^{n}\nu_k \cdot \eta_k(\forall\ n \in \mathrm{N}).$$

对此博弈模型,我们作下列假设:

假设 1 对弈过程独立同分布,即 $\eta=(\eta_n, n\in \mathrm{N})$ 独立,且 $P(\eta_n=1)=p, P(\eta_n=-1)=q=1-p$,其中 $0<p<1$.

假设 2 $\nu=(\nu_n, n\in N)$ 为可料序列.这个假设是基于甲方在第 $n+1$ 轮对弈时所下的赌注 ν_{n+1} 以此次之前的输出情况为依据这样一个现实背景.

假设 3 σ-代数流按如下格式给定:

$$\mathscr{F}_n = \sigma(\eta_0, \eta_1, \cdots \eta_n)\ (\forall\ n \in \mathrm{N}).$$

现在,我们将 ξ 写成鞅的 ν-变换形式.设 $Y_{-1}=0, Y_n=\sum_{k=0}^{n}\eta_k(n \in N)$,则 $\xi_n = \sum_{k=0}^{n}\nu_k \cdot \Delta Y_k(\forall\ n\in \mathrm{N})$.按假设 1~3,$Y$ 适应于 $(\mathscr{F}_\cdot)$,ν_{n+1} 是 $\mathscr{F}_n$-可测的.因此,ξ 是 Y 的 ν-变换.

若 $p=q$,则 $Y\in M(P,\mathscr{F}_\cdot)$,$\xi$ 为广义鞅.

若 $p>q$,则 $Y\in M_s(P,\mathscr{F}_\cdot)$,$\xi$ 为广义下鞅.

若 $p<q$,则 $Y\in M^s(P,\mathscr{F}_\cdot)$,$\xi$ 为广义上鞅.

上述结论是在假设 1~3 下得到的.这里利用了如下基本关系:

$$E(\Delta\xi_{n+1}/\mathscr{F}_n) = E(\nu_{n+1} \cdot \Delta Y_{n+1}/\mathscr{F}_n) = \nu_{n+1} \cdot (p-q).$$

不难看出，鞅表达平等对弈这一概念，而下鞅表达对弈于甲方有利. 上鞅表达对弈于甲方不利.

下面，考虑甲方在对弈中采取的一种对策. 如果甲方在直至 $n-1$ 轮对弈中都失败，则将第 n 轮的赌注加倍，而一旦成功就退出对弈. 这一对策的数学表示如下：$\nu_1=a$；当 $n\geqslant 2$ 时，

$$\nu_n = a \cdot 2^{n-1} \cdot 1_{\{\eta_k=-1,1\leqslant k\leqslant n-1\}}.$$

设 $\xi_0=0$，并令

$$\tau = \inf\{n \geqslant 1 : \eta_n = 1\}, \inf\{\varnothing\} = \infty.$$

显然，τ 为停时. $(\tau=m)$ 表示甲方直至 $m-1$ 轮都负，而在第 m 轮时胜. 因此，

$$\xi_n = \begin{cases} -\sum_{k=1}^{n}\nu_k = -a(2^n-1), & n < m, \\ a, & n = m. \end{cases} \quad 于(\tau = m)上.$$

终止时刻 τ 是一个很重要的变量. 现讨论它的某些性质.

性质 1　τ 为有限停时，即 $P(\tau<\infty)=1$.

事实上，$P(\tau=n)=P(\eta_k=-1,1\leqslant k\leqslant n-1;\eta_n=1)=q^{n-1}\cdot p$.

$$P(\tau < \infty) = \sum_{n=1}^{\infty}P(\tau = n) = p \cdot \sum_{n=1}^{\infty}q^{n-1} = \frac{p}{1-q} = 1.$$

性质 2　设 τ_p 为相应于 $p\in(0,1)$ 的停时，则

$$E\tau_p = \sum_{n=1}^{\infty}n \cdot P(\tau_p = n) = \sum_{n=1}^{\infty}n(q^{n-1} - q^n) = \frac{1}{1-q} = \frac{1}{p}.$$

（这个事实表明：当 $p\uparrow$ 时，甲方终止对弈时间就平均而言在缩短，这符合 p 值增高，对弈于甲方有利这一事实）.

性质 3　$\xi_\tau=a$(a.s)，且 $E\xi_\tau=a$.

（这个事实可解释为：若甲方有足够的赌资，而乙方一直陪赌，只要甲方赌技参数 $p>0$，他终归会达到目标值 a.）

性质 4　设 $a>0, p=\frac{1}{2}, \tau=\tau_p$，则 ξ 为鞅，且

$$E\xi_n = 0(\forall\ n \geqslant 1), E\xi_\tau = a(> 0).$$

(这种情形表明:当 n 用停时 τ 代替时,等式 $E\xi_n=0$ 不必保持. 但按引理 4.2.1,当 n 用停时 $n\wedge\tau$ 代替时,此等式必保持,即 $E\xi_{n\wedge\tau}=0$).

从性质 4 中很自然的提出了一个问题:当 ξ 为鞅时,在什么条件下,用停时 τ 代替 n 使得等式 $E\xi_\tau=E\xi_1$ 成立. 下一节将讨论这个问题.

§4.3 鞅的停时代替不变性

设 $\xi\in M(P,\mathscr{F}.)(M_s(P,\mathscr{F}.))$. τ 为停时. 本节要讨论的主要问题是:在什么条件下,当 n 由 τ 替代后,关系式 $E\xi_\tau=(\geqslant)E\xi_0$ 仍成立. 关于这个问题的最重要的结果就是 Doob 的停时代替不变性定理.

定理 4.3.1(Doob 停止定理) 设 $\xi\in M(P,\mathscr{F}.)(M_s(P,\mathscr{F}.))$. τ_1、τ_2 为有限停时. 若

(i) $E|\xi_{\tau_i}|<\infty\ (i=1、2)$;

(ii) $\lim\limits_{n\to\infty}\int_{(\tau_i>n)}|\xi_n|\mathrm{d}P=0\ (i=1,2)$,

则
$$E(\xi_{\tau_2}/\mathscr{F}_{\tau_1})=(\geqslant)\xi_{\tau_1}(\mathrm{mod}P)\text{ 于}(\tau_2\geqslant\tau_1)\text{ 上.}\tag{4.3-1}$$

特别地,若 $P(\tau_2\geqslant\tau_1)=1$,则
$$E(\xi_{\tau_2}/\mathscr{F}_{\tau_1})=(\geqslant)\xi_{\tau_1}(\mathrm{mod}P).\tag{4.3-2}$$

证 $\tau_1^*=\tau_2\wedge\tau_1$,则 τ_1^* 为停时,且 $\tau_1^*\leqslant\tau_2$. 这时,(4.3-1)式等价于:(4.3-2)式中将 τ_1 改成 τ_1^*. 因此,不失一般性,设 $P(\tau_2\geqslant\tau_1)=1$,并证明(4.3-2)式成立.

记 $A_n=A\cdot(\tau_1=n)\in\mathscr{F}_n\ (\forall A\in\mathscr{F}_{\tau_1})$. 按条件(i)及 Lebesque 控制收敛定理,有
$$E1_A\cdot(\xi_{\tau_2}-\xi_{\tau_1})=\lim_{N\to\infty}E[1_A\cdot(1_{(\tau_2\leqslant N)}\xi_{\tau_2}-1_{(\tau_1\leqslant N)}\xi_{\tau_1})].\tag{4.3-3}$$

$$E(1_A\cdot 1_{(\tau_1\leqslant N)}\xi_{\tau_1})=\sum_{n=0}^{N}E(1_{A_n\cdot(\tau_2\geqslant n)}\xi_n)$$

$$=\sum_{n=0}^{N}E(1_{A_n\cdot(\tau_2=n)}\xi_{\tau_2})+\sum_{n=0}^{N}E(1_{A_n\cdot(\tau_2\geqslant n+1)}\xi_n)$$

$$=(\leqslant)\sum_{n=0}^{N}E(1_{A_n\cdot(\tau_2=n)}\xi_{\tau_2})+\sum_{n=0}^{N}E(1_{A_n\cdot(\tau_2\geqslant n+1)}\xi_{n+1})$$

$$=(\leqslant)\sum_{n=0}^{N}E(1_{A_n\cdot(N\geqslant\tau_2\geqslant n)}\xi_{\tau_2})+\sum_{n=0}^{N}E(1_{A_n\cdot(\tau_2>N)}\xi_N)$$

$$=E(1_{A\cdot(\tau_2\leqslant N)}\xi_{\tau_2})+E(1_{A\cdot(\tau_1\leqslant N)\cdot(\tau_2>N)}\xi_N).$$

将此式代入(4.3-3)式中，得

$$E1_A.(\xi_{\tau_2}-\xi_{\tau_1})=(\geqslant)\varliminf_{N\to\infty}(-E(1_{A\cdot(\tau_1\leqslant N<\tau_2)}\xi_N))$$

按条件(ii)，有

$$\varliminf_{N\to\infty}|E(1_{(\tau_1\leqslant N<\tau_2)\cdot A}\xi_N)|\leqslant\varliminf_{N\to\infty}E(1_{(\tau_2>N)}|\xi_N|)=0.$$

故 $$E1_A\cdot(\xi_{\tau_2}-\xi_{\tau_1})=(\geqslant)0\ (\forall\ A\in\mathscr{F}_{\tau_1}).$$

此式等价于(4.3-2)式成立.

注 1 在上节的例 4.2.3 中构作了一个停时 τ，当 $p=\frac{1}{2}$，$a>0$ 时，有 $E\xi_\tau=a>0$，且

$$\lim_{m\to\infty}\int_{(\tau>m)}\xi_m\mathrm{d}P=a>0.$$

由此可见，此例中的鞅 ξ 不满足上述定理的条件(ii). 而且，实际上确实有 $E\xi_\tau=a\neq0=E\xi_1$.

注 2 如果条件(i)不成立，条件(ii)成立，也不能保证定理真. 设$(A_k,k\geqslant1)$为 Ω 的一个可测分割，且 $P(A_k)>0\ (\forall\ k\geqslant1)$. 让 $\mathscr{F}_n=\sigma(A_k,1\leqslant k\leqslant n)$，$\xi_n=\sum_{k=1}^{n}\frac{1}{kP(A_k)}1_{A_k}(n\geqslant1)$. 则 ξ 为下鞅. 定义映射 τ：$(\tau=n)=A_n$，则 τ 为停时，且

$$E\xi_\tau=\sum_{n=1}^{\infty}\frac{1}{n}=\infty,\int_{(\tau>m)}\xi_m\mathrm{d}P=0\ (\forall\ m\geqslant1).$$

可见下鞅 ξ 不满足上述定理中的条件(i). 虽对$\forall\ n\geqslant1$，有 $E(\xi_\tau/\mathscr{F}_n)\geqslant\xi_n$. 但它越出了可积范围.

推论 4.3.1 设$\xi\in M(P,\mathscr{F}.)$，且一致可积. τ、σ 为有限停时，且 $\tau\leqslant\sigma$(a.s)，则

$$E(\xi_\sigma/\mathscr{F}_\tau)=\xi_\tau(\text{a.s}),$$

证 为此，仅需验证定理 4.3.1 中的条件(i)和(ii).

ξ 为鞅表明：$|\xi|$为下鞅. 按引理 4.2.1，有

$$E|\xi_{n\wedge\tau}|\leqslant E|\xi_n|\ (\forall\ n\in\mathrm{N}).$$

ξ 一致可积表明：$\sup\limits_n E|\xi_n|<\infty$. 按 Fatou 引理，

$$\begin{aligned}E|\xi_\tau|&=E(\lim_{n\to\infty}|\xi_{n\wedge\tau}|)\leqslant\underline{\lim}_{n\to\infty}E|\xi_{n\wedge\tau}|\\&\leqslant\underline{\lim}_{n\to\infty}E|\xi_n|\leqslant\sup_n E|\xi_n|<\infty.\end{aligned}$$

同理，$E|\xi_\sigma|<\infty$. 故定理 4.3.1 中的条件(i)成立.

为证条件(ii)成立. 考虑如下估式：

$$\begin{aligned}\int_{(\tau>n)}|\xi_n|\mathrm{d}P&=\int_{\{|\xi_n|>c\}}|\xi_n|\cdot1_{(\tau>n)}\mathrm{d}P+\int_{\{|\xi_n|\leqslant c\}}|\xi_n|\cdot1_{(\tau>n)}\mathrm{d}P\\&\leqslant c\cdot P(\tau>n)+\sup_n\Big(\int_{\{|\xi_n|>c\}}|\xi_n|\mathrm{d}P\Big).\end{aligned}$$

按 ξ 一致可积，上式右边第二项当 $c\uparrow\infty$时极限为 0. 按 τ 的有限性，右边第一项当 $n\uparrow\infty$时极限为 0.

故
$$\underline{\lim}_{n\to\infty}\int_{(\tau>n)}|\xi_n|\mathrm{d}P=0.$$

同理，此式对 σ 亦成立，故条件(ii)满足.

推论 4.3.2 设$\xi\in A_s(\mathscr{F}.)$，且$\exists\ 0\leqslant\eta_\infty\in L^1$，使得

$$|\xi_n|\leqslant E(\eta_\infty/\mathscr{F}_n)\ (\text{a.s})\ (\forall\ n\in\mathrm{N}),$$

则对任意有限停时 τ，有

$$E|\xi_\tau|<\infty,\ \underline{\lim}_{n\to\infty}\int_{(\tau>n)}|\xi_n|\mathrm{d}P=0.$$

证 令 $\eta_n=E(\eta_\infty/\mathscr{F}_n)$. 如果 η 为一致可积鞅，则按推论 3.1，

$E\eta_\tau=E\eta_0=E\eta_\infty<\infty$，且$\underline{\lim}_{n\to\infty}\int_{(\tau>n)}\eta_n\mathrm{d}P=0$. 按$\xi$受$\eta$控制之假设，有$E|\xi_\tau|\leqslant E\eta_\tau<\infty$，

$$\underline{\lim}_{n\to\infty}\int_{(\tau>n)}|\xi_n|\mathrm{d}P\leqslant\underline{\lim}_{n\to\infty}\int_{(\tau>n)}\eta_n\mathrm{d}P=0.$$

剩下的问题就是证明：η为一致可积鞅.

η为鞅是显然的. 下证一致可积性.

$$P(\eta_n>C)\leqslant C^{-1}E\eta_n=C^{-1}E\eta_\infty$$

$$\Rightarrow\sup_n P(\eta_n>C)\leqslant C^{-1}E\eta_\infty\underset{C\to\infty}{\longrightarrow}0.$$

令$A_a=\{\eta_\infty\leqslant a\}$，$A_a^c=\{\eta_\infty>a\}$，则对$\forall\ a>0$，有

$$\sup_n\int_{\{\eta_n>C\}}\eta_n\mathrm{d}P=\sup_n\int_{\{\eta_n>C\}}\eta_\infty\mathrm{d}P$$

$$=\sup_n\left(\int_{A_a\{\eta_n>C\}}+\int_{A_a^c\{\eta_n>C\}}\right)\eta_\infty\mathrm{d}P$$

$$\leqslant a\sup_n P(\eta_n>C)+\int_{A_a^c}\eta_\infty\mathrm{d}P$$

$$\Rightarrow\lim_{C\to\infty}\sup_n\int_{\{\eta_n>C\}}\eta_n\mathrm{d}P=0.$$

注3　即使$\xi\in A_s(\mathscr{F}.)$一致可积，也不能保证：对$\forall$有限停时τ，有$E|\xi_\tau|<\infty$. 下例将说明这一点. 设$(A_n,n\geqslant1)$为Ω的一个可测分割，且$P(A_n)>0\ (\forall\ n\geqslant1)$，$\frac{1}{nP(A_n)}\uparrow\infty$当$n\uparrow\infty$时. $\mathscr{F}_n=\sigma(A_k:1\leqslant k\leqslant n)$. 定义

$$\xi_n=\frac{1}{nP(A_n)}1_{A_n}(n\geqslant1),$$

则$\xi\in A_s(\mathscr{F}.)$是一致可积的. 现在，构作停时τ，使得

$$(\tau=n)=A_n(\forall\ n\geqslant1).$$

对此停时，有$E|\xi_\tau|=\sum_{n=1}^{\infty}\frac{1}{n}=\infty$.

推论 4.3.3　设$\xi\in M(P,\mathscr{F}.)(M_s(P,\mathscr{F}.))$. τ为停时.

若$E\tau<\infty$，且$\exists\ n_0\in N$及常数C，使得

$$1_{(\tau\geqslant n)}E(|\Delta\xi_n|/\mathscr{F}_{n-1})\leqslant C\ (\text{a. s})\ (\forall\ n\geqslant n_0),$$

则 $E|\xi_\tau|<\infty$,且 $E\xi_\tau=(\geqslant)E\xi_0$.

证 设 $\eta_k=|\Delta\xi_k|,\xi_{-1}=0,\gamma_n=\sum_{k=0}^{n}\eta_k$,则

$$E\gamma_\tau=\sum_{n=0}^{\infty}E(1_{(\tau=n)}\sum_{k=0}^{n}\eta_k)=\sum_{k=0}^{\infty}\sum_{n=k}^{\infty}E(1_{(\tau=n)}\eta_k)$$

$$=\sum_{k=0}^{\infty}E(\eta_k\cdot 1_{(\tau\geqslant k)})\leqslant\sum_{k=0}^{n_0}E(\eta_k\cdot 1_{(\tau\geqslant k)})+C\sum_{k=n_0}^{\infty}P(\tau\geqslant k)$$

$$\leqslant\sum_{k=0}^{n_0}E\eta_k+C\cdot E\tau<\infty$$

$$\Rightarrow E|\xi_\tau|\leqslant E\gamma_\tau<\infty.$$

$$\varliminf_{n\to\infty}\int_{(\tau>n)}|\xi_n|\mathrm{d}P\leqslant\varliminf_{n\to\infty}(\int_{(\tau>n)}\gamma_n\mathrm{d}P)$$

$$\leqslant\varliminf_{n\to\infty}\int_{(\tau>n)}\gamma_\tau\mathrm{d}P=0.$$

故定理 4.3.1 中的条件(i)、(ii)成立.

定理 4.3.2(Wald 等式) 设 $\xi=(\xi_n,n\geqslant1)$为独立同分布的 r.v. 列. $\mathscr{F}_n=\sigma(\xi_1,\xi_2,\cdots,\xi_n)(n\geqslant1)$. τ 为停时,且 $E\tau<\infty$.

(i) 若 $\xi_1\in L^1$,则 $E\left(\sum_{k=1}^{\tau}\xi_k\right)=E\xi_1\cdot E\tau$

(ii) 若 $\xi_1\in L^2$,则 $E\left(\sum_{k=1}^{\tau}\xi_k-\tau E\xi_1\right)^2=E\tau V\xi_1$,.

其中 $V\xi_1=E(\xi_1-E\xi_1)^2$.

证 令 $\eta_n=\sum_{k=1}^{n}\xi_k-nE\xi_1=\sum_{k=1}^{n}(\xi_k-E\xi_1)$,则 $\eta\in M(P,\mathscr{F}.)$,

且 $$E(|\Delta\eta_n|/\mathscr{F}_{n-1})=E(|\xi_n-E\xi_1|/\mathscr{F}_{n-1})$$
$$=E|\xi_n-E\xi_1|\leqslant 2E|\xi_1|<\infty.$$

按推论 4.3.3,有 $E\eta_\tau=E\eta_1=0$. 由此即得(i).

为证(ii),补充定义 $\eta_0=0,\mathscr{F}_0=\{\varnothing,\Omega\}$. 则仍有 $\eta\in M(P,$

$\mathscr{F}.$)，其中$(\mathscr{F}.)=(\mathscr{F}_n,n\in\mathbb{N})$，且

$$0\leqslant E(\Delta\eta_n^2/\mathscr{F}_{n-1})=E((\Delta\eta_n)^2/\mathscr{F}_{n-1})$$
$$=E(\xi_n-E\xi_1)^2=V\xi_1<\infty.$$

由此即知：$\eta^2=(\eta_n^2,n\in\mathbb{N})$为非负下鞅，且

$$E\eta_\tau^2=\sum_{n=1}^{\infty}E(\eta_n^2 1_{(\tau=n)})=\sum_{n=1}^{\infty}E(1_{(\tau=n)}\sum_{n=1}^{\infty}\Delta\eta_n^2)$$
$$=\sum_{k=1}^{\infty}\sum_{n=k}^{\infty}E(\Delta\eta_k^2 1_{(\tau=n)})=\sum_{k=1}^{\infty}E(\Delta\eta_k^2\cdot 1_{(\tau\geqslant k)})$$
$$=\sum_{k=1}^{\infty}E(1_{(\tau\geqslant k)})E(\Delta\eta_k^2/\mathscr{F}_{k-1}))$$
$$=V\xi_1\sum_{k=1}^{\infty}P(\tau\geqslant k)=V\xi_1\cdot E\tau.$$

由此即得(ii).

注 4 如果 Wald 等式中，只假定 τ 为有限停时(不必可积)，则

$$\text{(i)}'\ \lim_{n\to\infty}\frac{1}{En\wedge\tau}E\Big(\sum_{k=1}^{n\wedge\tau}\xi_k\Big)=E\xi_1(\text{当 }\xi_1\in L^1\text{ 时});$$

$$\text{(ii)}'\ \lim_{n\to\infty}\frac{1}{En\wedge\tau}E\Big(\sum_{k=1}^{n\wedge\tau}\xi_k-(n\wedge\tau_k)E\xi_1\Big)^2=V\xi_1(\text{当 }\xi_1\in L^2\text{ 时}).$$

例 4.3.1 继续讨论例 4.2.3(博弈问题)，符号 $\eta=(\eta_n)_{n\geqslant 1}$，$\xi=(\xi_n)_{n\geqslant 1}$的含意不变. η 独立同分布.

$$P(\eta_n=1)=p,P(\eta_n=-1)=q=1-p,0<p<1.$$

假设甲方每次投赌一个单位，即 $\nu_k=1(\forall\ k\geqslant 1)$. 则甲方在前 n 轮对弈中输出的结果为 $\xi_n=\sum_{k=1}^{n}\eta_k(\forall\ n\geqslant 1)$. 设甲、乙分别有赌资$|A|$和 B，其中$-A$ 和 B 均为正整数. 如果甲、乙的赌资有一方输给对方，则对弈过程停止. 这可由如下停时表示：

$$\sigma=\inf\{n\geqslant 1:\xi_n=A\text{ 或 }B\},\inf\{\varnothing\}=\infty.$$

讨论的问题是：计算期望 $E\sigma$ 及毁灭概率

$$\alpha = P(\xi_\sigma = A), \beta = P(\xi_\sigma = B).$$

令 $\tau_n = n \wedge \sigma (n \geqslant 1)$，则 τ_n 为停时，且

(i) $\tau_n \uparrow \sigma$ 当 $n \uparrow \infty$ 时；

(ii) $|\xi_{\tau_n}| \leqslant |A| \vee B, \xi_{\tau_n}^2 \leqslant |A|^2 \vee B^2 (\forall\, n \geqslant 1)$.

若 $p \neq q$，则 $E\xi_1 \neq 0$. 按 Wald 第一等式，得

$$E\tau_n = \frac{1}{E\xi_1} E\xi_{\tau_n}$$

$$\Rightarrow E\sigma = \lim_{n\to\infty} E\tau_n = \frac{1}{E\xi_1} \lim_{n\to\infty} E\xi_{\tau_n} < \infty. \tag{4.3-4}$$

若 $p = q$，则 $E\xi_1 = 0, E\xi_1^2 = 1$. 按 Wald 第二等式，得

$$E\tau_n = E\xi_{\tau_n}^2$$

$$\Rightarrow E\sigma = \lim_{n\to\infty} E\tau_n = \lim_{n\to\infty} E\xi_{\tau_n}^2 < \infty. \tag{4.3-5}$$

由(4.3-4)式和(4.3-5)式可知：当 $0 < p < 1$ 时，$E\sigma < \infty$. 从而 σ 为有限停时，即 $P(\sigma < \infty) = 1$. 于是

$$\xi_{\tau_n} \xrightarrow{\text{a.s.}} \xi_\sigma \text{ 当 } n \to \infty \text{ 时}.$$

按 Lebesque 有界收敛定理，可推出

$$E\xi_\sigma = \lim_{n\to\infty} E\xi_{\tau_n}, E\xi_\sigma^2 = \lim_{n\to\infty} E\xi_{\tau_n}^2.$$

按(4.3-4)式和(4.3-5)式，得

$$E\sigma = \begin{cases} \dfrac{1}{p-q} E\xi_\sigma = \dfrac{1}{p-q}(A\alpha + B\beta), & \text{当 } p \neq q \text{ 时}, \\ E\xi_\sigma^2 = A^2\alpha + B^2\beta, & \text{当 } p = q \text{ 时}. \end{cases} \tag{4.3-6}$$

显然，由 $P(\sigma < \infty) = 1$ 可知

$$\alpha + \beta = 1. \tag{4.3-7}$$

为得到 α、β 之值，需要补充一个关系. 如何找到这个关系，我们将求助于 Doob 的鞅停时代替不变性定理 4.3.1.

若 $p = q$，则 $E\xi_1 = 0$，且 ξ 为鞅. 由 $\xi_\sigma^2 \in L'$ 及 σ 的有限性知：定理 3.1 中的条件满足. 于是，有

$$E\xi_\sigma = E\xi_1 = 0,\text{即 } A\alpha + B\beta = 0. \tag{4.3-8}$$

若 $p\neq q$，则 $E\xi_1\neq 0$，ξ 不为鞅. 这时，我们可以构造一个鞅 $\gamma=(\gamma_n)_{n\geqslant 1}$ 如下. 令

$$\gamma_n = \rho^{\xi_n}(n\geqslant 1),\text{其中 } \rho = \frac{q}{p}.$$

则 (i) $E|\gamma_n| = E\left(\prod_{k=1}^{n}\rho^{\eta_k}\right) = \prod_{k=1}^{n}E\rho^{\eta_k} = 1$;

(ii) $E(\gamma_{n+1}/\mathscr{F}_n) = \gamma_n\cdot E(\rho^{\eta_{n+1}}/\mathscr{F}_n) = \gamma_n(\text{a. s})$;

(iii) $\gamma_\sigma = \rho^{\xi_\sigma} \leqslant (\rho\vee\rho^{-1})^{|A|\vee B} < \infty$.

因此，γ 为鞅，且 σ 满足定理 4.3.1 中的要求. 于是

$$E\gamma_\sigma = E\gamma_1 = 1.\text{即 } \rho^A\alpha + \rho^B\beta = 1. \tag{4.3-9}$$

将(4.3-7)～(4.3-9)诸式联合起来，可解得

$$\alpha = \begin{cases} \dfrac{\rho^B - 1}{\rho^B - \rho^A}, & p\neq q, \\ \dfrac{B}{B+|A|}, & p = q \end{cases} \qquad \beta = \begin{cases} \dfrac{1-\rho_A}{\rho^B - \rho^A}, & p\neq q, \\ \dfrac{|A|}{B+|A|}, & p = q. \end{cases}$$

$$E\sigma = \begin{cases} \dfrac{1}{p-q}(A\alpha + B\beta), & p\neq q, \\ A^2\alpha + B^2\beta, & p = q. \end{cases}$$

其中 $0<p<1$，$\rho=\dfrac{q}{p}$.

注 此例中 α、β 的计算公式也可通过另一途径推导出来(参看[1]，ch. 1，§ 9)，简录如下：

补充：$\eta_0 = x\in(A,B)$. 令 $\xi_n^x = x+\xi_n(\forall\ n\geqslant 1)$，$\xi_0^x = x$.

$$\sigma_x = \inf\{n\in N: \xi_n^x = A \text{ 或 } B\},\inf\{\varnothing\} = \infty.$$

$$A_k^x = \bigcup_{i=0}^{k}\{\sigma_x = i,\xi_i^x = A\},B_k^x = \bigcup_{i=0}^{k}\{\sigma_x = i,\xi_i^x = B\}.$$

$$\alpha_k(x) = P(A_k^x),\beta_k(x) = P(B_k^x).$$

$$\alpha_k(x) + \beta_k(x) = P(\sigma_x\leqslant k).$$

由(i) $\alpha_k(x)\uparrow\alpha(x)$，$\beta_k(x)\uparrow\beta(x)$ ($\forall\ x\in[A,B]$为整数).

(ii) $\alpha_k(x) = E1_{A_k^x} = EE(1_{A_k^x}/\mathscr{F}_1) = pE1_{A_{k-1}^{x+1}} + q\cdot E1_{A_{k-1}^{x-1}}$

$=p\alpha_{k-1}(x+1)+q\alpha_{k-1}(x-1).$

(iii) $\alpha(x)=p\alpha(x+1)+q\alpha(x-1)$($\forall$ 整数 $x\in(A,B)$)

(iv) 边界条件:$\alpha(A)=1,\alpha(B)=0.$

可知,$\alpha(x)$,$\beta(x)$满足下列定解方程:

$$\begin{cases}\alpha(x)=p\alpha(x+1)+q\alpha(x-1)(\forall\ \text{整数}\ x\in(A,B)),\\ \alpha(A)=1,\alpha(B)=0.\end{cases} \tag{4.3-10}$$

$$\begin{cases}\beta(x)=p\beta(x+1)+q\beta(x-1)\ (\forall\ \text{整数}\ x\in(A,B)),\\ \beta(A)=0,\beta(B)=1.\end{cases} \tag{4.3-11}$$

分 $p=q$ 和 $p\neq q$ 两种情形解方程(4.3-10)和(4.3-11),即可得 $\alpha(x)$、$\beta(x)$的计算表示. 然后,让 $\alpha=\alpha(0)$,$\beta=\beta(0)$便可得毁灭概率 α 和 β,且 $\alpha+\beta=1$,从而,$P(\sigma<\infty)=1$.

令 $\tau_n(x)=n\wedge\sigma_x$,$m_n(x)=E\tau_n(x)$,则类似作法可推出:

$$m_n(x)=1+pm_{n-1}(x+1)+qm_{n-1}(x-1)$$
$$(\forall\ \text{整数}\ x\in(A,B)). \tag{4.3-12}$$

让 $n\uparrow\infty$,则得

$$\begin{cases}m(x)=pm(x+1)+qm(x-1)+1(\forall\ \text{整数}\ x\in(A,B))\\ m(A)=0,m(B)=0.\end{cases} \tag{4.3-13}$$

先在 $m(x)<\infty$的假设下求解方程(4.3-13). 然后,令 $E\sigma=m(0)$ 即可得所要的计算表达式. 然而,确定 $E\sigma$ 是否有限仍然要借助于 Wald 等式.

定理 4.3.3(Wald 基本方程) 设 $\xi=(\xi_n,n\geqslant1)$为独立同分布的 r.v. 列. $S_n=\sum\limits_{k=1}^{n}\xi_k(\forall\ n\geqslant1)$,$\mathscr{F}_n=\sigma(\xi_1,\xi_2,\cdots,\xi_n)$. 记 $\varphi(t)=E\exp(t\xi_1)(\forall\ t\in R)$. 并$\exists\ t_0\neq0$,$\ni 1\leqslant\varphi(t_0)<\infty$. 假如

(i) τ 为停时,且 $E\tau<\infty$;

(ii) $|S_n|\leqslant C$ 于$(\tau\geqslant n)$上($\forall\ n\geqslant n_0$),

其中 n_0 为某正整数. 则

$$E(\varphi^{-\tau}(t_0)\exp(t_0 S_\tau)) = 1.$$

证 令 $\eta_n=\varphi^{-n}(t_0)\exp(t_0 S_n)$, 则 $E\eta_n=1$, 且

$$E(\eta_{n+1}/\mathscr{F}_n) = \eta_n(\text{a.s}) \Rightarrow \eta \in M(P,\mathscr{F}_\cdot).$$

又

$$\begin{aligned}
&1_{(\tau\geqslant n)}E(|\Delta\eta_n|/\mathscr{F}_{n-1}) \\
&= 1_{(\tau\geqslant n)}E(\eta_{n-1}|\varphi^{-1}(t_0)\exp(t_0\xi_n)-1|/\mathscr{F}_{n-1}) \\
&= \eta_{n-1}1_{(\tau\geqslant n)}E(|\varphi^{-1}(t_0)\exp(t_0\xi_n)-1|) \leqslant 2\eta_{n-1}1_{(\tau\geqslant n)} \\
&\leqslant 2\exp(t_0|S_n|)1_{(\tau\geqslant n)} \leqslant 2\exp(t_0 C).
\end{aligned}$$

按推论 3.3, 得 $1=E(\varphi^{-1}(t_0)\cdot e^{t_0\xi_1})=E\eta_\tau$.

§ 4.4 基本不等式

本节主要介绍 Doob 的鞅不等式. 其它与鞅有关的不等式只给出结果, 其证明可参看有关文献. 定义下列符号:

$$\xi_n^* = \max_{0\leqslant k\leqslant n}|\xi_k|,\ \|\xi_n\|_p = (E|\xi_n|^p)^{\frac{1}{p}}(p>0).$$

定理 4.4.1(Doob 的鞅不等式) 设 $0\leqslant\xi\in M_s(P,\mathscr{F}_\cdot)$, 则对 $\forall\, n\in N$,

(i) $P(\xi_n^*\geqslant\varepsilon)\leqslant\frac{1}{\varepsilon}\int_{(\xi_n^*\geqslant\varepsilon)}\xi_n\mathrm{d}P\leqslant\frac{1}{\varepsilon}E\xi_n(\forall\ \varepsilon>0)$;

(ii) $\|\xi_n\|_p\leqslant\|\xi_n^*\|_p\leqslant\frac{p}{p-1}\|\xi_n\|_p(p>1)$;

(iii) $\|\xi_n\|_1\leqslant\|\xi_n^*\|_1\leqslant\frac{e}{e-1}[1+\|\xi_n\ln^+\xi_n\|_1]$.

证 先证(i). 设$(\mathscr{F}_\cdot)=(\mathscr{F}_k,0\leqslant k\leqslant n)$,

$$\tau_n = \inf\{0\leqslant k\leqslant n;\xi_k\geqslant\varepsilon\},\inf\{\varnothing\}=n.$$

则 τ_n 为有界停时, 按定理 4.3.1, $E(\xi_n/\mathscr{F}_{\tau_n})\geqslant\xi_{\tau_n}(\text{a.s})$

$$\begin{aligned}
\Rightarrow E\xi_n \geqslant E\xi_{\tau_n} &= \int_{(\xi_n^*\geqslant\varepsilon)}\xi_{\tau_n}\mathrm{d}P+\int_{(\xi_n^*<\varepsilon)}\xi_{\tau_n}\mathrm{d}P \\
&\geqslant \varepsilon P(\xi_n^*\geqslant\varepsilon)+E(\xi_n 1_{(\xi_n^*<\varepsilon)}).
\end{aligned}$$

于是结论(i)真.

次证(ii). 若 $\|\xi_n^*\|_p=0$ 或 $\|\xi_n\|_p=\infty$,则结论自然成立. 因此,仅需讨论 $0<\|\xi_n^*\|_p$ 及 $\|\xi_n\|_p<\infty$ 的情形. 按 ξ 的非负性及 Jensen 不等式,有

$$\|\xi_n^*\|_p^p \leqslant \sum_{k=0}^{n}\|\xi_k\|_p^p \leqslant (n+1)\|\xi_n\|_p^p < \infty.$$

由 $\xi_n \in L^p$ 可知

$$t^p \cdot P(\xi_n \geqslant t) \leqslant E(\xi_n^p 1_{\{\xi_n \geqslant t\}}) \to 0, \text{当 } t \to \infty \text{ 时}$$

$$\Rightarrow \|\xi_n^*\|_p^p = E|\xi_n^*|^p = p\int_0^\infty t^{p-1}P(\xi_n^* > t)\mathrm{d}t$$

$$\leqslant p\int_0^\infty t^{p-2}\int_{\{\xi_n^* \geqslant t\}}\xi_n \mathrm{d}P\mathrm{d}t$$

$$= pE\left(\int_0^\infty t^{p-2}1_{\{\xi_n^* \geqslant t\}}\mathrm{d}t \cdot \xi_n\right)$$

$$= \frac{p}{p-1}E(\xi_n \cdot (\xi_n^*)^{p-1}) \leqslant \frac{p}{p-1}\|\xi_n\|_p (E|\xi_n^*|^p)^{1-\frac{1}{p}}$$

$$\Rightarrow \|\xi_n^*\|_p \leqslant \frac{p}{p-1}\|\xi_n\|_p.$$

最后证(iii). 类似于上一步的计算可得

$$E\xi_n^* - 1 \leqslant E(\xi_n^* - 1)^+ = \int_0^\infty P(\xi_n^* - 1 > t)\mathrm{d}t$$

$$\leqslant \int_0^\infty \frac{1}{1+t}\int_{\{\xi_n^* \geqslant 1+t\}}\xi_n \mathrm{d}P\mathrm{d}t$$

$$\leqslant E\left(\xi_n\int_0^{(\xi_n^*-1)^+}\frac{\mathrm{d}t}{1+t}\right) = E(\xi_n \cdot \ln^+ \xi_n^*). \quad (4.4\text{-}1)$$

进一步的考虑需要用到如下不等式:

$$a\ln^+ b \leqslant a\ln^+ a + b\mathrm{e}^{-1} (\forall\, a \geqslant 0, b \geqslant 0). \quad (4.4\text{-}2)$$

假如此不等式成立,则令 $a=\xi_n, b=\xi_n^*$,便可得

$$E\xi_n^* - 1 \leqslant E(\xi_n \cdot \ln^+ \xi_n) + E(\mathrm{e}^{-1}\xi_n^*).$$

由此即得(iii). 下面证明(4.4-2)式.

令 $f(b)=a\ln^+ b - cb\ (c>0)$.

若 $a=0$ 或 $0\leqslant b\leqslant 1$，则 $f(b)=-cb$.

若 $0<a$，且 $1<b$，则 $f(b)=a\ln b-cb$，$f'(b)=\frac{a}{b}-c$.

因此，当 $a>c$ 时，$f'(b)=0$ 有解 $b=\frac{a}{c}$，且为 $f(b)$的极大值点，即 $f(b)\leqslant f\left(\frac{a}{c}\right)$. 当 $a\leqslant c$ 时，$f'(b)<0$. 从而，$f(b)\downarrow$. 这表明：$f(b)\leqslant f(1)=-c$.

综上分析得：

$$a\ln^+ b=f(b)+cb\leqslant\begin{cases}a\ln\left(\frac{a}{c}\right)-a+cb, \text{当 } a>0,b>1,0<c<a \text{ 时},\\ c(b-1)^+, \text{当 } 0\leqslant b\leqslant 1 \text{ 或 } a\leqslant c \text{ 时}.\end{cases}$$

在此式中，选 $c=e^{-1}$，即可推出(4.4.-2)式.

推论 4.4.1 设 $\xi\in M(P,\mathscr{F}.)$，且 $\xi_n\in L^2(\forall\ n\in N)$，则对$\forall\ n\in N$，有

(i) $P(\xi_n^*\geqslant\varepsilon)\leqslant\left(\frac{1}{\varepsilon}E|\xi_n|\right)\wedge\left(\frac{1}{\varepsilon^2}E|\xi_n|^2\right)(\forall\ \varepsilon>0)$；

(ii) $E|\xi_n^*|^2\leqslant 4E|\xi_n|^2$.

证 令 $\xi^2=(\xi_n^2,n\in N)$，则按 Jensen 不等式，$|\xi|$ 和 $\xi^2\in M_s(P,\mathscr{F}.)$. 按定理 4.4.1，命题中的(i)和(ii)成立.

定义 4.4.1 设 $\xi\in A_s(\mathscr{F}.)$，且非负. A 为适应增序列. 若对任意有限停时 τ，有 $E\xi_\tau\leqslant EA_\tau$，则称 ξ 受 A 控制.

定理 4.4.2 设 $\xi\in A_s(\mathscr{F}.)$，且非负，并受某可料增序列 A 的控制(A_0 非负). τ 为任意有限停时. 则下列不等式成立：

(i) $P(\xi_\tau^*\geqslant\varepsilon)\leqslant\frac{1}{\varepsilon}EA_\tau(\forall\ \varepsilon>0)$；

(ii) $P(\xi_\tau^*\geqslant\varepsilon)\leqslant\frac{1}{\varepsilon}E(a\wedge A_\tau)+P(A_\tau\geqslant a)\ (\forall\ \varepsilon>0,a>0)$；

(iii) $\|\xi_\tau^*\|_p\leqslant\left(\frac{2-p}{1-p}\right)^{\frac{1}{p}}\|A_\tau\|_p(0<p<1)$.

证 先证(i). 定义 $\sigma_n=\inf\{0\leqslant i\leqslant n\wedge\tau:\xi_i\geqslant\varepsilon\}$，$\inf\{\varnothing\}=n$，则 σ_n 为停时，$\sigma_n\uparrow\tau$(a. s.)，$A_{\sigma_n}\leqslant A_\tau$，$E\xi_{\sigma_n}\leqslant EA_{\sigma_n}$. 于是

$$\varepsilon P(\xi_{\sigma_n}^* \geqslant \varepsilon) \leqslant \int_{\{\xi_{\sigma_n}^* \geqslant \varepsilon\}} \xi_{\sigma_n} \mathrm{d}P \leqslant E\xi_{\sigma_n} \leqslant EA_{\sigma_n} \leqslant EA_\tau.$$

注意,$\{\xi_{\sigma_n}^* \geqslant \varepsilon\} \uparrow \{\xi_\tau^* \geqslant \varepsilon\}$. 由此及上式便可得(i).

(注:结论(i)的成立并不要求 A 可料.)

次证(ii). 定义 $\gamma = \inf\{k \in N: A_{k+1} \geqslant a\}$, $\inf\{\varnothing\} = \infty$,则由 A 的可料性知:γ 为停时,且 $A_\gamma < a$. 于是

$$\begin{aligned}
\{\xi_\tau^* \geqslant \varepsilon, A_\tau < a\} &= \{\xi_\tau^* \geqslant \varepsilon, A_\tau < a, \tau \leqslant \gamma\}; \\
&\quad \cup \{\xi_\tau^* \geqslant \varepsilon, A_\tau < a, \tau > \gamma\} \\
&= \{\xi_\tau^* \geqslant \varepsilon, A_\tau < a, \tau \leqslant \gamma\} = \{\xi_{\tau\wedge\gamma}^* \geqslant \varepsilon\} \cap \{\tau \leqslant \gamma\} \\
\Rightarrow P(\xi_\tau^* \geqslant \varepsilon) &= P(\xi_\tau^* \geqslant \varepsilon, A_\tau < a) + P(\xi_\tau^* \geqslant \varepsilon, A_\tau \geqslant a) \\
&= P(\xi_{\tau\wedge\gamma}^* \geqslant \varepsilon, \tau \leqslant \gamma) + P(\xi_\tau^* \geqslant \varepsilon, A_\tau \geqslant a) \\
&\leqslant P(\xi_{\tau\wedge\gamma}^* \geqslant \varepsilon) + P(A_\tau \geqslant a) \\
&\leqslant \frac{1}{\varepsilon} EA_{\tau\wedge\gamma} + P(A_\tau \geqslant a) \text{ (由结论(i))} \\
&\leqslant \frac{1}{\varepsilon} E(a \wedge A_\tau) + P(A_\tau \geqslant a).
\end{aligned}$$

最后证(iii). 按结论(ii),得

$$\begin{aligned}
\|\xi_\tau^*\|_p^p &= E|\xi_\tau^*|^p = \int_0^\infty P(|\xi_\tau^*|^p > t)\mathrm{d}t \\
&\leqslant \int_0^\infty t^{-\frac{1}{p}} E(t^{\frac{1}{p}} \wedge A_\tau)\mathrm{d}t + \int_0^\infty P(A_\tau^p \geqslant t)\mathrm{d}t \\
&= E\left(\int_0^{A_\tau^p} \mathrm{d}t\right) + E\left(\int_{A_\tau^p}^\infty A_\tau t^{-\frac{1}{p}} \mathrm{d}t\right) + EA_\tau^p \\
&= 2EA_\tau^p + \frac{p}{1-p} EA_\tau^p = \frac{2-p}{1-p} EA_\tau^p.
\end{aligned}$$

推论 4.4.2 设 $0 \leqslant \xi \in A_s(\mathscr{F}.)$,$A$ 为增序列(不必可料). $A_0 = 0$,且 ξ 受 A 的控制.

若存在常数 $c > 0$,于 $P(\sup_k \Delta A_k \leqslant C) = 1$. 则对任意有限停时 τ 及 $\forall\ \varepsilon > 0, a > 0$,有

$$P(\xi_\tau^* \geqslant \varepsilon) \leqslant \frac{1}{\varepsilon} E((a + c) \wedge A_\tau) + P(A_\tau \geqslant a).$$

证 定义 $\gamma=\inf\{k\in N: A_k\geqslant a\}$, $\inf\{\varnothing\}=\infty$. 则 γ 为停时, $A_{\gamma-1}<a$, $A_{-1}=0$(规定), $A_\gamma=A_\gamma-A_{\gamma-1}+A_{\gamma-1}\leqslant C+a$(a.s), $A_{\tau\wedge\gamma}\leqslant A_\gamma\leqslant C+a$ (a.s), $A_{\tau\wedge\gamma}\leqslant(C+a)\wedge A_\tau$(a.s). 于是

$$\{\xi_\tau^*\geqslant\varepsilon, A_\tau<a\}=\{\xi_\tau^*\geqslant\varepsilon,\tau<\gamma\}=\{\xi_{\tau\wedge\gamma}^*\geqslant\varepsilon,\tau<\gamma\}$$

$$\begin{aligned}\Rightarrow P(\xi_\tau^*\geqslant\varepsilon)&=P(\xi_\tau^*\geqslant\varepsilon,A_\tau<a)+P(\xi_\tau^*\geqslant\varepsilon,A_\tau\geqslant a)\\&\leqslant P(\xi_{\tau\wedge r}^*\geqslant\varepsilon)+P(A_\tau\geqslant a)\\&\leqslant\frac{1}{\varepsilon}EA_{\tau\wedge\gamma}+P(A_\tau\geqslant a)\\&\leqslant\frac{1}{\varepsilon}E((a+c)\wedge A_\tau)+P(A_\tau\geqslant a).\end{aligned}$$

下列关于鞅的不等式之证可参看文献[9].

定理 4.4.3(Khinchin 不等式) 设 $\xi=(\xi_n,n\geqslant1)$为独立同分布的 Bernoulli 序列,且 $P(\xi_1=1)=\frac{1}{2}=P(\xi_1=-1)$. $C=(C_n,n\geqslant 1)$为任意实数列,则对 $\forall\ 0<p<\infty$, $\exists\ A_p>0$, $B_p>0$, ∋

$$A_p\Big(\sum_{k=1}^n C_k^2\Big)^{\frac{1}{2}}\leqslant\Big\|\sum_{k=1}^n C_k\xi_k\Big\|_p\leqslant B_p\Big(\sum_{k=1}^n C_k^2\Big)^{\frac{1}{2}}(\forall\ n\geqslant1).$$

定理 4.4.4(Marcinkiewicz-Zygmund 不等式). 设 $\xi=(\xi_n,n\geqslant1)$为独立的 r.v. 列,且 $E\xi_n=0(\forall\ n\geqslant1)$. 则对$\forall\ p\geqslant1$, $\exists$ 常数 A_p,B_p,使得

$$A_p\Big\|\Big(\sum_{k=1}^n\xi_k^2\Big)^{\frac{1}{2}}\Big\|_p\leqslant\Big\|\sum_{k=1}^n\xi_k\Big\|_p\leqslant B_p\Big\|\Big(\sum_{k=1}^n\xi_k^2\Big)^{\frac{1}{2}}\Big\|_p.$$

注 上述两不等式的共同特点是:被估量的变量均构成一个自生鞅. 这就导致考虑更一般的情形.

定义 4.4.2 若$\sup\limits_n\|\xi_n\|_p<\infty$,则称 ξ 是 L^p-有界的.

定理 4.4.5(Burkholder 不等式) 设 $\xi\in M(P,\mathscr{F}.)$,且 L^p-有界,其中 $p>1$, $\xi_0=0$. 则$\exists\ A_p(=B_p^{-1})$, $B_p\left(=\frac{18p^{3/2}}{(p-1)^{1/2}}\right)$, ∋$A_p\Big\|\Big(\sum_{k=1}^n(\Delta\xi_k)^2\Big)^{\frac{1}{2}}\Big\|_p\leqslant\|\xi_n\|_p\leqslant B_p\Big\|\Big(\sum_{k=1}^n(\Delta\xi_k)^2\Big)^{\frac{1}{2}}\Big\|_p(\forall\ n$

$\geqslant 1$),

$$A_p \| \Big(\sum_{k=1}^{\infty} (\Delta\xi_k)^2 \Big)^{\frac{1}{2}} \|_p \leqslant \sup_n \| \xi_n \|_p \leqslant B_p \| \Big(\sum_{k=1}^{\infty} (\Delta\xi_k)^2 \Big)^{\frac{1}{2}} \|_p.$$

(注. B.-不等式中要求 $p>1$. 而下面的 D.-不等式则考虑 $p=1$ 的情形.)

定理 4.4.6(Davis 不等式)　设 $\xi \in M(\mathrm{P}, \mathscr{F}.)$,且 L'-有界,则$\exists$ $0<A<B$,使得

$$AE\Big(\sum_{k=1}^{\infty} (\Delta\xi_k)^2 \Big)^{\frac{1}{2}} \leqslant E(\sup_n |\xi_n|) \leqslant BE(\sum_{k=1}^{\infty} (\Delta\xi_k)^2)^{\frac{1}{2}}.$$

§4.5　半鞅的收敛性

本节讨论收敛性问题. 这里,下(或上)穿数的估计将起着特别重要的作用. 这种估计基于这样一个事实:"一个实数列 $x=(x_n, n \in \mathrm{N})$ 具有极限(含 $\pm\infty$)"等价于"此数列 x 在任意两个有理数 a、b $(a<b)$ 之间的振动次数是有限的."

设 $\xi \in A_s(\mathscr{F}.)$,即适应于($\mathscr{F}.$)的 r. v. 列,让 $a, b \in \mathrm{R}$,且 $a<b$, $\tau_0=0$, $\tau_{2m-1}=\inf\{k>\tau_{2(m-1)}: \xi_k \leqslant a\}$,

$$\tau_{2m} = \inf\{k > \tau_{2m-1}: \xi_k \geqslant b\}, \inf\{\varnothing\} = \infty.$$

显然,τ_k 为停时,且 $\tau_k \uparrow \infty$ 当 $k \uparrow \infty$时.

定义 4.5.1　令 $\beta_n(a,b)=\begin{cases} \max\{m \geqslant 1: \tau_{2m} \leqslant n\}, & \tau_2 \leqslant n, \\ 0, & \tau_2 > n. \end{cases}$

$\xi^n=(\xi_k, 0 \leqslant k \leqslant n)$,则称 $\beta_n(a,b)$ 为 ξ^n 关于区间 (a,b) 的上穿数. 令

$$\beta(a,b) = \lim_{n \to \infty} \beta_n(a,b),$$

则称 $\beta(a,b)$ 为 ξ 关于区间 (a,b) 的上穿数.

显然,上穿数 $\beta(a,b)$ 就是 ξ 从下面的 a 通过上面的 b 的次数. 从 β_n, β 的定义中可以看出:它们都是取值于 N 的 r. v. 当 $\beta(a,b)=\infty$ 时,$\xi(\omega)$ 一定不收敛. 现在,来估计这个随机数.

定理 4.5.1(Doob 上穿不等式)　设 $\xi \in M_s(P, \mathscr{F}.)$,则对$\forall$ n

$\geqslant 1, a, b \in R\ (a<b)$，有

$$E\beta_n(a,b) \leqslant \frac{1}{b-a}E(\xi_n - a)^+.$$

证 设 $\eta=\frac{1}{b-a}(\xi-a)^+$，则 $\eta \in M_s(P,\mathscr{F}.)$，且

$$\beta'_n(0,1) = \beta_n(a,b).$$

其中 $\beta'_n(0,1)$ 表示 η 关于 $(0,1)$ 的上穿数. 于是，问题转为证明：$E\beta'_n(0,1) \leqslant E\eta_n$. 因此，不妨设 $a=0, b=1, \xi$ 非负. 定义 $\sigma_k = n \wedge \tau_k$，$A_k = \{(i,\omega): \sigma_k < i \leqslant \sigma_{k+1}\}$，$\varphi_i^{(k)}(\omega) = \nu_k 1_{A_k}(i,\omega)$，其中 $\nu_k = \begin{cases}1, k \text{ 为奇数},\\ 0, k \text{ 为偶数}.\end{cases}$ A_k^i 表示 A_k 的 i-截集. 显然，$A_k^i \in \mathscr{F}_{i-1}$. 因此，$\varphi_i^{(k)}$ 是 $\mathscr{F}_{i-1}$-可测的 r.v.

按上述定义，有下列各式成立：

$$\nu_k(\xi_{\sigma_{k+1}} - \xi_{\sigma_k}) = \sum_{i=0}^{n} \varphi_i^{(k)} \Delta\xi_i;$$

$$\nu_k(\xi_{\tau_{k+1}} - \xi_{\tau_k}) \geqslant \nu_k \text{ 于} (\tau_{k+1} < \infty) \text{ 上};$$

$$\sum_{k=0}^{\infty} \varphi_i^{(k)} \leqslant 1 (\forall\, i \in N).$$

$$\beta_n(0,1) = \sum_{k=0}^{\infty} \nu_k \cdot 1_{(\tau_{k+1} \leqslant n)} \leqslant \sum_{k=0}^{\infty} \nu_k \cdot (\xi_{\sigma_{k+1}} - \xi_{\sigma_k}) = \sum_{k=0}^{\infty} \sum_{i=0}^{n} \varphi_i^{(k)} \Delta\xi_i$$

$$\Rightarrow E\beta_n(0,1) \leqslant \sum_{k=0}^{\infty} \sum_{i=0}^{n} E(\varphi_i^{(k)} E(\Delta\xi_i / \mathscr{F}_{i-1}))$$

$$= \sum_{i=1}^{n} E(\sum_{k=0}^{\infty} \varphi_i^{(k)} E(\Delta\xi_i / \mathscr{F}_{i-1}))$$

$$\leqslant \sum_{i=1}^{n} EE(\Delta\xi_i / \mathscr{F}_{i-1}) = \sum_{i=1}^{n} E\Delta\xi_i = E(\xi_n - \xi_0) \leqslant E\xi_n.$$

注 1 按 Fatou 引理及 $\beta=(\beta_n, n\geqslant 1)$ 的单调性，有

$$E\beta(a,b) \leqslant \varliminf_{n\to\infty} \frac{1}{b-a}E(\xi_n - a)^+ = \frac{1}{b-a}\sup_n E(\xi_n - a)^+.$$

定理 4.5.2（Doob 鞅收敛定理） 设 $\xi \in M_s(P,\mathscr{F}.)$，且 L^1-有

界. 则 $(a.s)-\lim\limits_{n\to\infty}\xi_n=\xi_\infty\in L^1$.

证 假如 ξ 不是 a.s-收敛的. 则必有

$$P(\overline{\lim}\xi_n > \underline{\lim}\xi_n) > 0.$$

这表明: $\exists\, a,b\in R, a<b$, 使得

$$P(\overline{\lim}\xi_n \geqslant b > a \geqslant \underline{\lim}\xi_n) > 0, \text{即 } P(\beta(a,b)=\infty) > 0.$$

另一方面, 按注 5.1 中的估式, 有

$$E\beta(a,b) \leqslant \underline{\lim}_{n\to\infty}\frac{1}{b-a}E(\xi_n-a)^+$$

$$\leqslant\frac{1}{b-a}(|a|+\sup_n E|\xi_n|) < \infty.$$

从而, $\beta(a,b)<\infty$(a.s). 这导致矛盾. 故 ξ(a.s)收敛.

按 Fatou 引理, 有

$$E|\xi_\infty| \leqslant \underline{\lim}E|\xi_n| \leqslant \sup_n E|\xi_n| < \infty \Rightarrow \xi_\infty \in L^1.$$

注 2 定理中的"ξ 是 L^1-有界的"可以改成"ξ^+ 是 L^1-有界的". 事实上, $\xi\in M_s(P,\mathscr{F}.)$ 表明:

$$\sup_n E\xi_n^+ \leqslant \sup_n E|\xi_n| = \sup_n(2E\xi_n^+ - E\xi_n)$$

$$\leqslant \sup_n(2E\xi_n^+ - E\xi_0) \leqslant 2\sup_n E\xi_n^+ + E|\xi_0|.$$

推论 4.5.1 若 $\xi\in M(P,\mathscr{F}.)$, 且 L^1-有界, 则

$$(a.s)-\lim_{n\to\infty}\xi_n=\xi_\infty\in L^1.$$

证 $M(P,\mathscr{F}.)\subseteq M_s(P,\mathscr{F}.)$ 表明: 此结论为定理 4.5.2 的特例.

推论 4.5.2 若 $\xi\in M^s(P,\mathscr{F}.)$, 且非负. $\bar\xi=(\xi_n,n\in\overline{\mathrm{N}})$, 这里 $\mathscr{F}_\infty=\sigma(\bigcup_n\mathscr{F}_n)$, $\xi_\infty=\overline{\lim\limits_{n\to\infty}}\xi_n$. $(\overline{\mathscr{F}.})=(\mathscr{F}_n,n\in\overline{\mathrm{N}})$, 则 $\bar\xi\in M^s(P,\overline{\mathscr{F}.})$, 即非负上鞅 ξ 在补充元素 $(\xi_\infty,\mathscr{F}_\infty)$ 之后仍为非负上鞅.

证 显然, $-\xi\in M_s(P,\mathscr{F}.)$, 且 $(-\xi)^+=0$ 是 L^1-有界的. 按注 2, 有 $(a.s)-\lim\limits_{n\to\infty}\xi_n=\xi_\infty\in L^1$. 按 Fatou 引理及 $\xi\in M^s(P,\mathscr{F}.)$,

$$E(\xi_\infty/\mathscr{F}_n) \leqslant \underline{\lim}_{m\to\infty}E(\xi_m/\mathscr{F}_n) \leqslant \xi_n(a.s).$$

故 $\bar{\xi}\in M^{s}(P,\overline{\mathscr{F}}.)$.

注3 在定理 4.5.2 中,即使 ξ 为鞅,且 L^1-有界也只能得到 ξ (a.s)收敛于 $\xi_\infty\in L^1$,不能保证收敛按 L^1 成立. 原因在于这些条件不能保证 ξ 一致可积. 例如,设 ξ 为独立同分布的 Bernoulli r.v. 列,且 $P(\xi_0=0)=\frac{1}{2}=P(\xi_0=1)$. 令 $\eta_n=\prod_{i=0}^{n}(2\xi_i)$,$\mathscr{F}_n=\sigma(\xi_i,0\leqslant i\leqslant n)$,则 η 为一个 L^1-有界的非负鞅. 然而,η 不是一致可积的. 事实上,对$\forall$ $C>0$,当 $2^{n+1}>C$ 时,有

$$\int_{\{\eta_n>C\}}\eta_n \mathrm{d}P = 2^{n+1}P(\eta_n = 2^{n+1}) = 2^{n+1}\cdot P^{n+1}(\xi_0 = 1) = 1.$$

这表明 η 不满足一致可积性要求.

推论 4.5.3 若 $\xi\in M(P,\mathscr{F}.)$,且一致可积,则

$$\bar{\xi} \in M(P,\mathscr{F}.),且\ \xi_n = E(\xi_\infty/\mathscr{F}_n)\ (\text{a.s})(\forall\ n \in \mathrm{N}).$$

同时 ξ(a.s)和 L^1-收敛于 ξ_∞(一般称此类鞅为右闭鞅或 Doob 鞅).

证 ξ 为一致可积鞅表明:ξ 为 L^1-有界鞅. 按推论 4.5.1,(a.s)$-\lim\limits_{n\to\infty}\xi_n=\xi_\infty\in L^1$,现在,令 $\xi'_n=\xi_n-\xi_\infty$,则 $\xi'_n\to 0$(a.s). 按 Doob 不等式及$|\xi|\in M_s(P,\mathscr{F}.)$,有

$$P(\sup_n|\xi_n| \geqslant C) = \lim_{n\to\infty}P(\sup_{0\leqslant k\leqslant n}|\xi_k| \geqslant C) \leqslant \frac{1}{C}\lim_{n\to\infty}E|\xi_n|$$

$$\leqslant \frac{1}{C}\sup_n E|\xi_n| \xrightarrow[c\to\infty]{} 0.$$

于是,由 ξ 一致可积及 $\xi_\infty\in L^1$ 可得

$$E(|\xi'_n|1_{\{|\xi_n|>C\}}) \leqslant E(|\xi_n|1_{\{|\xi_n|>C\}}) + E(|\xi_\infty|1_{\{\sup_n|\xi_n|>C\}})$$

$$\leqslant \sup_n E(|\xi_n|1_{\{|\xi_n|>C\}}) + E(|\xi_\infty|1_{\{\sup_n|\xi_n|>C\}}) \xrightarrow[c\to\infty]{} 0.$$

按 Lebesque 控制收敛定理,有

$$\lim_{n\to\infty}E(|\xi'_n|1_{\{|\xi_n|\leqslant C\}}) = 0(\forall\ C > 0)$$

$$\Rightarrow \lim_{n\to\infty}E|\xi'_n| = 0,即\ \xi_n \xrightarrow{L^1} \xi_\infty.$$

由 $\xi\in M(P,\mathscr{F}.)$可知:

$$E|E(\xi_\infty/\mathscr{F}_n)-\xi_n|\leqslant E|\xi_\infty-\xi_m|\xrightarrow[m\to\infty]{}0(\forall\ n\in\mathrm{N})$$

$$\Rightarrow\xi_n=E(\xi_\infty/\mathscr{F}_n)\ (\mathrm{a.s})(\forall\ n\in\mathrm{N}).$$

定理 4.5.3(鞅表现定理) $\xi\in A_s(\mathscr{F}.)$为一致可积鞅的充要条件是

(i) $\exists\ \eta\in L^1,\ni\xi_n=E(\eta/\mathscr{F}_n)\ (\mathrm{a.s})\ (\forall\ n\in\mathrm{N})$;

(ii) $\overline{\xi}$ 为鞅.

证 显然,条件(i)和(ii)等价.条件的必要性就是推论 4.5.3 的结论.下证充分性.不妨设条件(i)成立.

显然,ξ 为鞅,$|\xi|$为下鞅.按 Doob 的鞅不等式,有

$$P(\sup_n|\xi_n|\geqslant C)=\lim_{n\to\infty}P(\max_{0\leqslant k\leqslant n}|\xi_k|\geqslant C)$$

$$\leqslant\frac{1}{C}\sup_n E|\xi_n|\leqslant\frac{1}{C}E|\eta|\xrightarrow[c\to\infty]{}0$$

$$\Rightarrow\int_{\{|\xi_n|>C\}}|\xi_n|\mathrm{d}P\leqslant\int_{\{|\xi_n|>C\}}|\eta|\mathrm{d}P\leqslant\int_{\{\sup\limits_n|\xi_n|>C\}}|\eta|\mathrm{d}P\xrightarrow[c\to\infty]{}.$$

故ξ 一致可积.

推论 4.5.4 设$\xi=(\xi_n,n\in\mathrm{N})$一致可积,且$\xi_n\xrightarrow{L^1}\xi_\infty$,则

$$L^1\text{-}\lim_{n,m\to\infty}E(\xi_m/\mathscr{F}_n)=E(\xi_\infty/\mathscr{F}\infty).$$

证 令

$$\psi_{m,n}=E(\xi_m/\mathscr{F}_n),\psi_{\infty,\infty}=E(\xi_\infty/\mathscr{F}\infty),\psi_{\infty,n}=E(\xi_\infty/\mathscr{F}_n),$$

则 $$E|\psi_{m,n}-\psi_{\infty,\infty}|\leqslant E|\psi_{m,n}-\psi_{\infty,n}|+E|\psi_{\infty,n}-\psi_{\infty,\infty}|$$

$$\leqslant E|\xi_m-\xi_\infty|+E|\psi_{\infty,n}-\psi_{\infty,\infty}|.$$

按假设,此式右边第一项趋于 0.按推论 4.5.3,第二项趋于 0.故 $\lim\limits_{m,n\to\infty}E|\psi_{m,n}-\psi_{\infty,\infty}|=0$.

注 4 在上述命题中,极限与路线无关.因此,取 $m=n$,则所得到的结果就是 Doob-Levy 定理.

定理 4.5.4 设$\xi\in M_s(P,\mathscr{F}.)$,且$L^p$-有界(某 $p>1$),则$\overline{\xi}\in M_s(P,\overline{\mathscr{F}}.)$,其中$\xi_\infty=\overline{\lim\limits_{n\to\infty}}\xi_n^+$,$(\overline{\mathscr{F}}.)=(\mathscr{F}_n,n\in\overline{\mathrm{N}})$.

证 显然，ξ^+ 为 L'-有界下鞅，且一致可积. 按Doob 的鞅不等式，

$$\|\sup_k(\xi_k^+)\|_p = \lim_{n\to\infty}\|\sup_{0\leqslant k\leqslant n}(\xi_k^+)\|_p$$

$$\leqslant \frac{p}{p-1}\sup_n\|\xi_n^+\|_p < \infty.$$

按定理 4.5.2，(a.s) $-\lim_{n\to\infty}\xi_n^+ = \xi_\infty^+ \in L'$. 按 Lebesque 控制收敛定理，$L'-\lim_{n\to\infty}\xi_n^+ = \xi_\infty^+$. 按推论 4.5.4，有

$$E(\xi_\infty^+/\mathscr{F}_n) = \lim_{m\to\infty}E(\xi_m^+/\mathscr{F}_n) \geqslant \xi_n^+ \geqslant \xi_n(\text{a.s})\ (\forall\ n \in \mathrm{N});$$

故 $\overline{\xi} \in M_s(P,\overline{\mathscr{F}}.)$.

注 5 下鞅 ξ 右闭的更好的充分条件是：$\sup_n\xi_n^+ \in L'$.

定理 4.5.5 设 $\xi \in M(P,\mathscr{F}.)$（或 $M_s(P,\mathscr{F}.)$），且一致可积（非负且$\sup_n\xi_n \in L'$）. τ_i 为停时$(i=1,2)$，且 $\tau_1 \leqslant \tau_2$(a.s).

$\xi_\infty = \overline{\lim_{n\to\infty}}\xi_n$，则 $\xi^* = (\xi_{\tau_i}, \mathscr{F}_{\tau_i}, i=1,2)$ 为二元鞅（下鞅）.

证 按推论 4.5.3（注 5），有

$$E(\xi_\infty/\mathscr{F}_n) = \xi_n(\geqslant \xi_n)(\text{a.s})(\forall\ n \in \mathrm{N}),$$

即
$$\overline{\xi} \in M(P,\overline{\mathscr{F}}.)(M_s(P,\overline{\mathscr{F}}.)).$$

于是
$$E(|\xi_\infty|/\mathscr{F}_n) \geqslant |\xi_n|(\text{a.s})(\forall\ n \in \mathrm{N}).$$

因此

$$E(|\xi_{\tau_i}|) = \sum_{n=0}^{\infty}E(|\xi_n|1_{(\tau_i=n)}) + E(|\xi_\infty|1_{(\tau_i=\infty)})$$

$$\leqslant \sum_{n=0}^{\infty}E(E(|\xi_\infty|/\mathscr{F}_n)1_{(\tau_i=n)}) + E(|\xi_\infty|1_{(\tau_i=\infty)}).$$

$$= E|\xi_\infty| < \infty\ (i = 1、2).$$

按推论 4.5.4 及有界停时代换的不变性，有

$$E(\xi_{\tau_2}/\mathscr{F}_{\tau_1}) = \lim_{n\to\infty}E(\xi_{n\wedge\tau_2}/\mathscr{F}_{n\wedge\tau_1})$$

$$= (\geqslant)\lim_{n\to\infty}\xi_{n\wedge\tau_1} = \xi_{\tau_1}(\text{a.s}).$$

故 ξ^* 为二元鞅（下鞅）.

推论 4.5.5 设 $\xi=(\xi_n,n\geqslant 1)$ 为独立的 r. v. 列，$S_n=\sum_{k=0}^{n}\xi_k(\forall\ n\in N)$，则 $S=(S_n,n\in \mathrm{N})$ 的下列三种收敛意义是等价的：

(i) (a.s)收敛；(ii) P-收敛；(iii) d-收敛.

证 在第二章中已得结论：(i)⇒(ii)⇒(iii). 此关系一般不可逆. 然而，对本命题则是可逆的. 为证此，仅需证明：(ii)⇒(i).

假设 $S_n\xrightarrow{d}S_\infty$，即

$$E\exp(itS_n)\xrightarrow[n\to\infty]{}E\exp(itS_\infty)(\forall\ t\in \mathrm{R}).$$

显然，$\exists\ \delta>0$，⊣当 $|t|<\delta$ 时，$|E\exp(itS_\infty)|>0$. 于是，对 $\forall\ |t_0|<\delta$，$\exists\ C_0>0$ 及 $n_0\in N$，⊣当 $n\geqslant n_0$ 时，$|E\exp(it_0S_n)|\geqslant C_0$.

令
$$\mathscr{F}_n=\sigma(\xi_k,0\leqslant k\leqslant n)\ (n\in N),$$

$$\eta_n=\frac{\exp(it_0S_n)}{E\exp(it_0S_n)}\ (n\geqslant n_0),$$

$\eta_n=E(\eta_{n0}/\mathscr{F}_n)(0\leqslant n\leqslant n_0)$. 则 $\eta\in M(P,\mathscr{F}.)$，且

$$\sup_n E|\eta_n|\leqslant \sup_{n\geqslant n_0}|E\exp(it_0S_n)|^{-1}\leqslant\frac{1}{C_0}<\infty,$$

即 η 是 L'-有界的. 于是，按定理 4.5.2，有

$$(\mathrm{a.s})\text{-}\lim_{n\to\infty}\eta_n=\eta_\infty\in L'\Rightarrow\exp(it_0S_n)\xrightarrow[n\to\infty]{\mathrm{a.s}}\exp(it_0S_\infty).$$

现在，令 $I_\delta=\{t:|t|<\delta\}$，$\mathscr{B}_\delta=I_\delta\cap\mathscr{B}$. λ 表示 $(I_\delta,\mathscr{B}_\delta)$ 上的 Lebesque 测度. 定义

$$\mathscr{U}=\{(t,\omega)\in I_\delta\times\Omega:\lim_{n\to\infty}\exp(itS_n)\ 存在\},$$

$$\mathscr{U}^t=\{\omega:(t,\omega)\in\mathscr{U}\},\mathscr{U}^\omega=\{t:(t,\omega)\in\mathscr{U}\}.$$

则 $\mathscr{U}\in\mathscr{B}_\delta\times\mathscr{F}$，$P(\mathscr{U}^t)=1$. 按 Fubini 定理，有

$$\int_{I_\delta\times\Omega}1_{\mathscr{U}(t,\omega)}\mathrm{d}(\lambda\times P)=\int_{I\delta}\left(\int_\Omega 1_{\mathscr{U}}(t,\omega)P(\mathrm{d}\omega)\right)\lambda(\mathrm{d}t)$$

$$=\int_{I\delta}P(\mathscr{U}^t)\lambda(\mathrm{d}t)=\lambda(I_\delta).$$

另一方面

$$\lambda(I_\delta)=\int_\Omega\int_{I\delta}1_{\mathscr{U}}(t,\omega)\lambda(\mathrm{d}t)P(\mathrm{d}\omega)=\int_\Omega\lambda(\mathscr{U}^\omega)P(\mathrm{d}\omega).$$

显然，$\lambda(I_\delta)\geqslant\lambda(\mathscr{U}^\omega)(\forall\ \omega\in\Omega)$. 于是，$\exists\ \tilde{\Omega}\subset\Omega$，使得

$$P(\tilde{\Omega})=1,\text{且}\ \lambda(\mathscr{U}^\omega)=\lambda(I_\delta)(\forall\ \omega\in\tilde{\Omega}).$$

这表明：当 $t\in\mathscr{U}^\omega,\omega\in\tilde{\Omega}$ 时，$\lim\limits_{n\to\infty}\exp(itS_n)$存在. 按下面的引理，即可得：$S_n(\omega)\to S_\infty(\omega)\ (\forall\ \omega\in\tilde{\Omega})$，即 $S_n\xrightarrow{\text{a.s}}S_\infty$.

引理 4.5.1 设 $a_n,n\geqslant1$，为实数列. 若$\exists\ \delta>0$，$\ni$当$|t|<\delta$时，极限$\lim\limits_{n\to\infty}\exp(ita_n)$存在，则极限$\lim\limits_{n\to\infty}a_n$ 存在且有限，即$\lim\limits_{n\to\infty}a_n=a\in R$.

证 令 $f(t)=\lim\limits_{n\to\infty}\exp(ita_n)(|t|<\delta)$，则易证：$\exists\ a\in\mathrm{R}$，$\ni f(t)=\exp(ita)(|t|<\delta)$. 假如 $a,a_n(n\geqslant1)$以 M 为界. 注意，$\lim\limits_{n\to\infty}\exp(it(a_n-a))=1(|t|<\delta)$. 因此，

$$\lim_{n\to\infty}\sin(t(a_n-a))=0\ (|t|<\delta).$$

让 $t_0=\frac{1}{2}\left(\frac{1}{M}\wedge\delta\right)$，则$\lim\limits_{n\to\infty}\sin(t_0(a_n-a))=0$. 从而$\lim\limits_{n\to\infty}a_n=a$.

现在，证明，$(a_n,n\geqslant1)$无界是不可能的. 若不然，令 $\bar{a}_n=a_n-a$，则$(\bar{a}_n,n\geqslant1)$无界. 不失一般性，设 $\bar{a}_n\geqslant0$，且 $\bar{a}_n\uparrow\infty$. 给定 $t_0\in(0,\delta)$，则有分解 $t_0\bar{a}_n=k_n(t_0)\pi+r_n(t_0)$，其中 $k_n(t_0)$为非负整数，$k_n(t_0)\uparrow\infty$，$-\frac{\pi}{2}\leqslant r_n(t_0)<\frac{\pi}{2}$. 于是

$$\lim_{n\to\infty}\sin r_n(t_0)=\lim_{n\to\infty}\sin(k_n(t_0)\pi+r_n(t_0))=0.$$

由此即得：$\lim\limits_{n\to\infty}r_n(t_0)=0$.

若$(k_n(t_0),n\geqslant1)$中有无穷多个奇整数. 不妨设它全为奇数(否则选子列). 注意

$$0=\lim_{n\to\infty}\sin(t\bar{a}_n)=\lim_{n\to\infty}\sin\left(\frac{t}{t_0}t_0\bar{a}_n\right)$$

$$=\lim_{n\to\infty}\sin\left(\frac{t}{t_0}K_n(t_0)\pi+\frac{t}{t_0}r_n(t_0)\right)$$

$$=\lim_{n\to\infty}\sin\left(\frac{t}{t_0}K_n(t_0)\pi\right).\ (\forall\ 0<|t|<\delta).$$

选取 $t=\frac{1}{2}t_0$，则$\lim_{n\to\infty}\sin\left(\frac{t}{t_0}K_n(t_0)\pi\right)=1$ 或$-1\neq0$. 这不可能. 故 $(K_n(t_0), n\geqslant1)$中不可能有无穷多个奇数，假如$(K_n(t_0), n\geqslant1)$中有无穷多个偶数，不妨设它全为偶数，则取 $t=\frac{t_0}{4}$，即可知 $\left(\frac{1}{2}K_n(t_0), n\geqslant1\right)$ 中不可能有无穷多个奇数. 按此推论即知，$(K_n(t_0), n\geqslant1)$只可能有无穷多个形如 $K_n(t_0)=2^{m_n}$ 的偶数，其中 m_n 为整数，且 $m_n\uparrow\infty$当 $n\uparrow\infty$时，易证：有分解$\frac{1}{3}2^{m_n}=u_n+e_n$，其中 u_n 为非负整数，$e_n=\frac{1}{3}$ 或$\frac{2}{3}$. 选取 $t=\frac{t_0}{3}$，则 $0=\lim_{n\to\infty}\sin\left(\frac{1}{3}K_n(t_0)\pi\right)=\lim_{n\to\infty}\sin(e_n\pi)\neq0$. 此矛盾表明，$(K_n(t_0), n\geqslant1)$不可能有无穷多个偶数. 故$(K_n(t_0), n\geqslant1)$只能有有限个整数. 故$(\bar{a}_n, n\geqslant1)$不能无界.

§4.6 广义半鞅的收敛性

本节主要讨论上节的结果到广义半鞅的推广. 符号 $GM(P,\mathscr{F}.)$, $GM_s(P,\mathscr{F}.)$, $GM^s(P,\mathscr{F}.)$分别表示广义鞅、广义下鞅、广义上鞅类. 假如 ξ 为广义半鞅，按定义 4.2.1(ii)，它应满足要求：

$$E(\xi_{n+1}^{+}/\mathscr{F}_n)\wedge E(\xi_{n+1}^{-}/\mathscr{F}_n)<\infty(\text{a.s})\ (\forall n\in N).$$

定理 4.6.1 设 $\xi\in GM_s(P,\mathscr{F}.)$，$\beta_{n,m}(a,b)$为$(\xi_n,\cdots,\xi_m)$关于$(a,b)$的上穿数$(a<b)$，则

$$E(\beta_{n,m}(a,b)/\mathscr{F}_n)\leqslant\frac{1}{b-a}E((\xi_m-a)^{+}/\mathscr{F}_n)(\text{a.s})(a<b, n\leqslant m).$$

证 类似于定理 4.5.1，可将问题转化为证明如下不等式：$E(\beta_{n,m}(0,1)/\mathscr{F}_n)\leqslant E(\xi_m/\mathscr{F}_n)(\text{a.s})(n\leqslant m)$，其中 ξ 非负.

设 $A\in\mathscr{F}_n$，$\eta_k=1_A\xi_{k+n}$，则 η 为广义下鞅，且

$$\beta'_{m-n}(0,1)=1_A\beta_{n,m}(0,1),$$

其中 $\beta'_{m-n}(0,1)$ 为 $(\eta_0,\eta_1,\cdots,\eta_{m-n})$ 关于 $(0,1)$ 的上穿数. 按定理 4.5.1，$E\beta'_{m-n}(0,1)\leqslant E\eta_{m-n}$，即

$$E(\beta_{n,m}(0,1)\cdot 1_A)\leqslant E(\xi_m\cdot 1_A)\ (\forall\ A\in\mathscr{F}_n).$$

推论 4.6.1 设非负 $\xi\in GM^s(P,\mathscr{F}.)$，则对 $\forall\ a<b$，有

$$E(\beta_{n,m}(a,b)/\mathscr{F}_n)\leqslant\frac{1}{b-a}b^+\ (\text{a. s})(n\leqslant m).$$

证 显然，$-\xi\in GM_s(P,\mathscr{F}.)$，$\overline{\beta}_{n,m}(\cdot,\cdot)$ 表示 $-\xi$ 的 (n,m) 节的上穿数，则

$$\overline{\beta}_{n,m}(-b,-a)=\beta_{n,m}(a,b)(\forall\ a<b).$$

按定理 4.6.1，得

$$E(\overline{\beta}_{n,m}(-b,-a)/\mathscr{F}_n)\leqslant\frac{1}{b-a}E((-\xi_m-(-b))^+$$

$$/\mathscr{F}_n)\leqslant\frac{b^+}{b-a}.$$

定理 4.6.2 设 $\xi\in GM_s(P,\mathscr{F}.)$，则

$$A=\{\omega:\inf_k\sup_m E(\xi_m^+/\mathscr{F}_k)<\infty\}.$$

则 $\xi_\infty=(\text{a. s})-\lim\limits_{n\to\infty}\xi_n$ 存在于集 A 上，且有限或 $-\infty$.

证 设 $\xi_k^n=(\xi_k-n)^+$，则 $\xi^n\in GM_s(P,\mathscr{F}.)$，且非负. 令

$$A_{r_1r_2}^n=\{\omega:\overline{\lim_{k\to\infty}}\xi_k^n\geqslant r_2>r_1\geqslant\underline{\lim_{k\to\infty}}\xi_k^n\}\in\mathscr{F}_\infty(r_2>r_1),$$

$K_1=\bigcup\limits_{n=-\infty}^{0}\bigcup\{A_{r_2r_1}^n:r_1、r_2\in Q,r_1<r_2\}$，$Q$ 为有理数全体. 则 K_1 为 ξ 无极限存在之集.

首先证明：$\inf\limits_k E(1_{A_{r_1r_2}^n}/\mathscr{F}_k)>0$ 于 $A_{r_1r_2}^n$ 上(a. s).

事实上，若 $P(A_{r_1r_2}^n)=0$，则此结论自然成立. 现设 $P(A_{r_1r_2}^n)>0$ 并令 $\eta_k=E(1_{A_{r_1r_2}^n}/\mathscr{F}_k)$，则 η 为一致可积鞅，且非负. 定义停时

$$\tau=\inf\{k\in N:\eta_k=0\},\inf\{\varnothing\}=\infty,$$

则 $$0 < P(A^n_{r_1r_2}) = EE(1_{A^n_{r_1r_2}}/\mathscr{F}_\tau)$$

$$= \sum_{k=0}^{\infty} E(\eta_k 1_{(\tau=k)}) + E(\eta_\infty 1_{(\tau=\infty)})$$

$$= P(A^n_{r_1r_2} \cap (\tau = \infty))$$

$$\Rightarrow A^n_{r_1r_2} \subseteq (\tau = \infty)\ (\mathrm{mod}P).$$

按 τ 之定义，$\eta_k>0$ 于 $(\tau=\infty)$ 上 $(\forall\ k\in \mathrm{N})$. 因此，$\eta_k>0$ 于 $A^n_{r_1r_2}$ 上 $(\forall\ k\in N)$. 由 η 的一致可积性及 $A^n_{r_1r_2}\in\mathscr{F}_\infty$ 可知：$\eta_k \xrightarrow{\text{a. s}} \eta_\infty = 1_{A^n_{r_1r_2}}$. 故 $\inf\limits_k E(1_{A^n_{r_1r_2}}/\mathscr{F}_k)>0$ 于 $A^n_{r_1r_2}$ 上(a. s).

次证 $\inf\limits_k \sup\limits_m E(\xi_m^+/\mathscr{F}_k)=\infty$ 于 K_1 上(a. s).

事实上，$\beta_{k,\infty}(r_1,r_2)=\lim\limits_{m\to\infty}\beta_{k,m}(r_1,r_2)=\infty$ 于 $A^n_{r_1r_2}$ 上(a. s).

按定理 4. 6. 1，有

$$E(\beta_{k,\infty}(r_1,r_2)/\mathscr{F}_k) \leqslant \frac{1}{r_2-r_1}(|r_1| + \sup_{m\geqslant k} \cdot E(\xi_m^+/\mathscr{F}_k))\ (\forall\ k\in \mathrm{N}). \quad (4.6\text{-}1)$$

而 $$\inf_k E(\beta_{k,\infty}(r_1,r_2)/\mathscr{F}_k) \geqslant \infty \cdot \inf_k E(1_{A^n_{r_1r_2}}/\mathscr{F}_k)$$

$$=\infty \text{ 于 } A^n_{r_1r_2} \text{ 上(a. s)}.$$

于是， $$\inf_k \sup_m E(\xi_m^+/\mathscr{F}_k) = \infty \text{ 于 } A^n_{r_1r_2} \text{ 上(a. s)}$$

$$\Rightarrow \inf_k \sup_m E(\xi_m^+/\mathscr{F}_k) = \infty \text{ 于 } K_1 \text{ 上(a. s)}.$$

最后证明：$K_2\subseteq K_3(\mathrm{mod}P)$，其中

$$K_2 = \{\omega: \lim_{k\to\infty}\xi_k = +\infty\},\ K_3 = \{\omega: \inf_k \sup_m E(\xi_m^+/\mathscr{F}_k) = \infty\}.$$

为此，令

$$W_u = \{\omega: \sup_m E(\xi_m^+/\mathscr{F}_k) < u\} \in \mathscr{F}_k (\forall\ u>0).$$

若有某 $u>0$，$\ni P(W_u\cap K_2)>0$，则

$$\sup_m \int_{W_u} E(\xi_m^+/\mathscr{F}_k)\mathrm{d}P = \sup_m \int_{W_u} \xi_m^+ \mathrm{d}P \geqslant \sup_m \int_{W_u\cap K_2} \xi_m^+ \mathrm{d}P$$

$$\geqslant \int_{W_u\cap K_2} \lim_{m\to\infty} \xi_m^+ \mathrm{d}P = \infty \cdot P(W_u \cap K_2) = \infty;$$

$$\sup_m\int_{W_u}E(\xi_m^+/\mathscr{F}_k)\mathrm{d}P\leqslant\int_{W_u}\sup_m E(\xi_m^+$$
$$/\mathscr{F}_k)\mathrm{d}P\leqslant u\cdot P(W_u)<\infty.$$
$$\Rightarrow P(W_u\cap K_2)=0\ (\forall\ u>0)$$
$$\Rightarrow K_2\subseteq K_3(\mathrm{mod}P).$$

综上分析，得

$$K_1\cup K_2\subset K_3(\mathrm{mod}P)$$
$$\Rightarrow A=K_3^c\subseteq(K_1\cup K_2)^c.$$

注1 若$\xi\in GM_s(P,\mathscr{F}.)$，且$\xi\in A_s(\mathscr{F}.)$，则$A=A'$，其中$A=\{\omega:\inf\limits_k\sup\limits_m E(\xi_m^+/\mathscr{F}_k)<\infty\}$，$A'=\{\omega:\inf\limits_k\sup\limits_{m\geqslant k}E(\xi_m^+/\mathscr{F}_k)<\infty\}$. 证明此结论并不困难，但若$\xi$为广义的适应序列，则一般只能有$A\subseteq A'$. 然而，审查上述定理之证明立即可以发现，定理关于A'亦成立.

定理 4.6.3(Levy 定理的推广) 设$\bar{\eta}$为广义 r.v. 且$E\bar{\eta}>-\infty$，则

(i) $\xi_\infty=(a.s)-\lim\limits_{n\to\infty}\xi_n$存在(含取$\pm\infty$)，其中$\xi_n=E(\bar{\eta}/\mathscr{F}_n)$；

(ii) $E(\bar{\eta}/\mathscr{F}_\infty)\leqslant\xi_\infty(\mathrm{a.s})$；

(iii) $\xi_\infty=E(\bar{\eta}/\mathscr{F}_\infty)$ (a.s)当$P(\xi_\infty<\infty)=1$时.

证 先证(i). 令$\eta_n^\pm=E(\bar{\eta}^\pm/\mathscr{F}_n)$，则$\eta^-$为一致可积鞅. 按定理 4.5.2 和 4.5.3，有

$$\eta_n^-\xrightarrow{\mathrm{a.s}}\eta_\infty^-=E(\bar{\eta}^-/\mathscr{F}_\infty).$$

按注 1，$\eta_\infty^+=(\mathrm{a.s})-\lim\limits_{n\to\infty}\eta_n^+$存在，且有限于集$A'$上，这里$A'=\{\omega:\inf E(\bar{\eta}^+/\mathscr{F}_k)<\infty\}$. 若$E(\bar{\eta}^+/\mathscr{F}_n)=\infty(\forall\ n\in N)$，则自然有$\lim\limits_{n\to\infty}\eta_n^+=\infty$. 故$(\mathrm{a.s})-\lim\limits_{n\to\infty}\eta_n^+$存在(含取$+\infty$). 综合起来得

$$\xi_\infty=\eta_\infty^+-\eta_\infty^-=(a.s)-\lim_{n\to\infty}\xi_n.$$

次证(ii). 对$\forall\ C>0$，按(Levy)定理 4.5.3，有

$$E(\bar{\eta}^+\wedge C/\mathscr{F}_\infty)=(\mathrm{a.s})-\lim_{n\to\infty}E(\bar{\eta}^+\wedge C/\mathscr{F}_n)\leqslant\eta_\infty^+.$$

按 Fatou 引理，$E(\bar{\eta}^+/\mathscr{F}_\infty)\leqslant\eta_\infty^+$(a.s)，故

$$E(\bar{\eta}/\mathscr{F}_\infty)\leqslant\xi_\infty(\text{a.s}).$$

最后证(iii). 按(i)中之证，有 $\eta_\infty^-=E(\bar{\eta}^-/\mathscr{F}_\infty)$ (a.s). 因此，仅需证明 $\eta_\infty^+=E(\bar{\eta}^+/\mathscr{F}_\infty)$ (a.s). 按假设，有 $P(\eta_\infty^+<\infty)=1$. 令

$$A_C=\{\eta_\infty^+<C\},A'_{C,k}=\bigcup_{n=k}^{\infty}\{\eta_n^+<C\},A_{C,k}=\bigcap_{n=k}^{\infty}\{\eta_n^+<C\},$$

则 $A'_{C,k}\downarrow A_C, A_{C,k}\uparrow A_C, A_C\uparrow\Omega$ 当 $k\uparrow\infty, C\uparrow\infty$ 时，易证 $E(\bar{\eta}^+/\mathscr{F}_\infty)1_{\{\eta_k^+<C\}}$ 是可积的($\forall\, k\in N$). 按定理 4.5.3，有

$$\begin{aligned}
E(\bar{\eta}^+/\mathscr{F}_\infty)1_{A'_{C,k}} &\geqslant E(\bar{\eta}^+/\mathscr{F}_\infty)1_{\{\eta_k^+<C\}}\\
&=\lim_{n\to\infty}E(E(\bar{\eta}^+/\mathscr{F}_\infty)1_{\{\eta_k^+<C\}}/\mathscr{F}_n)\\
&=1_{\{\eta_k^+<C\}}\lim_{n\to\infty}E(\bar{\eta}^+/\mathscr{F}_n)=1_{\{\eta_k^+<C\}}\eta_\infty^+.\ (\text{a.s})\\
&\geqslant\eta_\infty^+1_{A_{C,k}}(\forall\, k\in N)\\
&\Rightarrow E(\bar{\eta}^+/\mathscr{F}_\infty)1_{A_C}=\lim_{k\to\infty}(E(\bar{\eta}^+/\mathscr{F}_\infty)1_{A'_{C,k}})\\
&\geqslant\lim_{k\to\infty}(\eta_\infty^+1_{A_{C,k}})=\eta_\infty^+1_{AC}.
\end{aligned}$$

利用本定理中的结论(ii)，便可得

$$E(\bar{\eta}^+/\mathscr{F}_\infty)=\eta_\infty^+ \text{ 于 } A_C(\text{a.s})\ (\forall\, C>0)$$
$$\Rightarrow E(\bar{\eta}^+/\mathscr{F}_\infty)=\eta_\infty^+(\text{a.s}).$$

定理 4.6.4 设 $\xi\in GM(P,\mathscr{F}.)$，且$\exists\ \mu\in L'$，使得

$$\xi_n\geqslant E(\mu/\mathscr{F}_n)\ (\text{a.s})\ (\forall\, n\in N),$$

则(i) $\xi_\infty=(\text{a.s})-\lim\limits_{n\to\infty}\xi_n$ 存在.

(ii) ξ_∞在集$\{\xi_0<\infty\}$上(a.s)有限；

(ii) $E(\xi_\infty/\mathscr{F}_0)\leqslant\xi_0$(a.s).

证 令 $\eta_n=\xi_n-E(\mu/\mathscr{F}_n)$，则 $\eta\in GM(P,\mathscr{F}.)$，且非负. 因此，按定理 4.6.3，$\eta_\infty=(\text{a.s})-\lim\limits_{n\to\infty}\eta_n$ 存在. 按定理 4.5.3，$\xi_\infty=(\text{a.s})-\lim\limits_{n\to\infty}\xi_n$ 存在. 按定理 4.6.2，η_∞于集 A 上有限或$-\infty$，其中 $A=\{\omega:\inf\limits_k\sup\limits_m E(\eta_m^+/\mathscr{F}_k)<\infty\}$. η_∞非负. 因此，η_∞于集 A 上(a.s)有

限. 考虑到:

$$\inf_k \sup_m E(\eta_m^+/\mathscr{F}_k) = \inf_k \max_{0\leqslant m\leqslant k} \eta_m = \eta_0.$$

$$\Rightarrow A = \{\eta_0 < \infty\} = \{\xi_0 < \infty\}.$$

从而, ξ_∞ 在集 $\{\xi_0<\infty\}$ 上(a.s)有限. 最后按 Fatou 引理, 得

$$E(\eta_\infty/\mathscr{F}_0) \leqslant \underline{\lim} E(\eta_n/\mathscr{F}_0) = \eta_0 (\text{a.s}).$$

$$\Rightarrow E(\xi_\infty/\mathscr{F}_0) \leqslant \xi_0 (\text{a.s}).$$

注 2 按注 1, 可以用 A' 代替 A:

$$A' = \{\omega: \inf_k \lim_{m\to\infty} E(\eta_m^+/\mathscr{F}_k) < \infty\} = \{\omega: \inf_k \eta_k^+ < \infty\}.$$

显然, $\{\omega: \eta_n^+<\infty\}\subset A'$ ($\forall\, n\in N$). 于是, ξ_∞ 在 $\bigcup_k \{\xi_k<\infty\}$ 上(a.s)有限. 因此, 只要有某个 ξ_k 为 r.v. 则 ξ_∞ 为 r.v.

定理 4.6.5 设 $\xi\in GM_s(P,\mathscr{F}.)$, 且 $E(\sup_n \xi_n^+)<\infty$. τ,σ 为停时, 且 $\tau\leqslant\sigma$. $\xi_\infty=\overline{\lim_{n\to\infty}}\xi_n$. 则 $E(\xi_\sigma/\mathscr{F}_\tau)\geqslant\xi_\tau$(a.s).

证 按假设, $\sup_n \xi_n^+ \in L'$. 因此,

$$\inf_k \sup_m E(\xi_m^+/\mathscr{F}_k) \leqslant \inf_k E(\sup_m \xi_m^+ /\mathscr{F}_k) \leqslant E(\sup_m \xi_m^+/\mathscr{F}_0) < \infty \ (\text{a.s}).$$

按定理 4.6.2, $\xi_\infty=(\text{a.s})-\lim_{n\to\infty}\xi_n$ 存在, 且有限或 $-\infty$.

令 $\xi_n^C=\xi_n\vee C (C<0)$, 则 ξ^C 为右闭下鞅. 按定理 4.5.5, $E(\xi_\sigma^C/\mathscr{F}_\tau)\geqslant\xi_\tau^C$(a.s). 按 Fatou 引理, 有

$$\overline{\lim_{C\to-\infty}} \xi_\tau^C \leqslant \overline{\lim_{C\to-\infty}} E(\xi_\sigma^C/\mathscr{F}_\tau) \leqslant E(\overline{\lim_{C\to-\infty}} \xi_\sigma^C/\mathscr{F}_\tau). \quad (4.6\text{-}2)$$

按 ξ_τ^C 关于 $C\downarrow-\infty$ 的单调性, 有

$$\overline{\lim_{C\to-\infty}} \xi_\tau^C = \lim_{C\to-\infty} \xi_\tau^C = \lim_{C\to-\infty} (\xi_\tau^C 1_{(\tau<\infty)} + \xi_\infty^C 1_{(\tau=\infty)})$$

$$= \xi_\tau 1_{(\tau<\infty)} + \lim_{C\to-\infty} \lim_{n\to\infty} \xi_n^C (1_{\{\xi_\infty=-\infty\}} + 1_{\{\xi_\infty>-\infty\}}) 1_{(\tau=\infty)}$$

$$= \xi_\tau 1_{(\tau<\infty)} + 1_{(\tau=\infty)} (1_{\{\xi_\infty=-\infty\}}\cdot(-\infty) + 1_{(\xi_\infty>-\infty)}\xi_\infty) = \xi_\tau.$$

同理可证 $\overline{\lim_{C\to-\infty}} \xi_\sigma^C=\xi_\sigma$. 代入(4.6-2)式即得结论.

注 3 广义上鞅反号即为广义下鞅. 因此, 调整一下符号便可

得到广义上鞅的相应结果.

§4.7 上鞅分解和势

假如 $\xi\in M^s(P,\mathscr{F}.)$，且 $\exists\ \eta\in L^1$，$\ni \xi_n\geq E(\eta/\mathscr{F}_n)=\eta'_n$ (a.s) $(\forall\ n\in\mathrm{N})$. 令 $\xi'_n=\xi_n-\eta'_n(\forall\ n\in\mathrm{N})$. 则 $\xi_n=\xi'_n+\eta'_n(\forall\ n\in\mathrm{N})$，这里 $\xi'\in M^s(P,\mathscr{F}.)$，且非负，$\eta'$ 为 Doob-Levy 鞅. 由此可见，ξ 可以分解成非负上鞅与鞅之和. 但这种分解未涉及到它们的极限行为. 从§5中可以看出，它们的极限与它们的结构有密切关系. 本节将从结构的角度来讨论上鞅的分解问题.

定理 4.7.1（Riesz 分解） 设 $\xi\in M^s(P,\mathscr{F}.)$，且 $\sup\limits_n E\xi_n^-<\infty$，则 ξ 可唯一地分解成如下形式：

$$\xi_n=\zeta_n+\eta_n(\forall\ n\in\mathrm{N}),$$

其中 $\eta\in M(P,\mathscr{F}.)$，$\zeta\in M^s(P,\mathscr{F}.)$，且 ζ 非负，$E\zeta_n\xrightarrow[n\to\infty]{}0$.

证 令 $\gamma_{k,n}=E(\xi_{k+n}/\mathscr{F}_k)$，则由 ξ 的上鞅性知

$$\begin{aligned}\gamma_{k,n}&=E(E(\xi_{k+n}/\mathscr{F}_{k+n-1})/\mathscr{F}_k)\leqslant E(\xi_{k+n-1}/\mathscr{F}_k)\\&=\gamma_{k,n-1}\leqslant\xi_k(\text{a.s}).\end{aligned}$$

因此，对 $\forall\ k\in\mathrm{N}$，$\gamma_{k,n}$ 关于 $n\in\mathrm{N}$ 是单调不增的，且上控. 令 $\eta_k=(\text{a.s})-\lim\limits_{n\to\infty}\gamma_{k,n}$，则 η_k 是 $\mathscr{F}_k$-可测的. 按单调收敛定理，$E\eta_k=\lim\limits_{n\to\infty}E\gamma_{k,n}\leqslant E\xi_k<\infty$. 按假设，有

$$-\infty<-\sup_k E\xi_k^-\leqslant\lim_{n\to\infty}E\xi_{k+n}=\lim_{n\to\infty}E\gamma_{k,n}=E\eta_k.$$

故 $\eta_k\in L^1(\forall\ k\in\mathrm{N})$. 再一次利用单调收敛定理，得

$$\begin{aligned}\eta_k&=\lim_{n\to\infty}\gamma_{k,n}=\lim_{n\to\infty}E(\gamma_{k+1,n-1}/\mathscr{F}_k)\\&=E(\lim_{n\to\infty}\gamma_{k+1,n-1}/\mathscr{F}_k)=E(\eta_{k+1}/\mathscr{F}_k)\ (\text{a.s}).\end{aligned}$$

综上分析，得 $\eta\in M(P,\mathscr{F}.)$.

现在，令 $\zeta_n=\xi_n-\eta_n(\forall\ n\in\mathrm{N})$，则 ζ 为非负上鞅，且 $E\zeta_n\to0$. 事实上，按 ξ 为上鞅，η 为鞅及单调收敛定理，下列各式成立：

$$\zeta_n = \xi_n - \lim_{m\to\infty}\gamma_{n,m}$$
$$= \lim_{m\to\infty}(\xi_n - E(\xi_{n+m}/\mathscr{F}_n)) \geqslant 0;$$
$$\zeta_n \geqslant E(\xi_{n+1}/\mathscr{F}_n) - E(\eta_{n+1}/\mathscr{F}_n)$$
$$= E(\xi_{n+1} - \eta_{n+1}/\mathscr{F}_n) = E(\zeta_{n+1}/\mathscr{F}_n);$$
$$E\zeta_n = E\xi_n - E\eta_n = E\xi_n - \lim_{m\to\infty}E\gamma_{n,m}$$
$$= E\xi_n - \lim_{m\to\infty}E\xi_m \xrightarrow[n\to\infty]{} 0.$$

故ζ满足定理中的要求.

最后证明唯一性. 假定有二个同格式的分解:

$$\zeta'_n + \eta'_n = \xi_n = \zeta''_n + \eta''_n (\forall\ n\in \mathrm{N}),$$

则 $\eta_n \triangleq \eta'_n - \eta''_n = \zeta''_n - \zeta'_n \triangleq \zeta_n$. 考虑到: $\eta\in M(P,\mathscr{F})$, $E|\zeta_n| \leqslant E|\zeta'_n| + E|\zeta''_n|$. 有 $E|\zeta_n|\to 0$. 从而, $E|\eta_n|\to 0$. 注意, $|\eta|$为非负下鞅. 于是, $E|\eta_n| \leqslant E|\eta_{n+m}| \xrightarrow[m\to\infty]{} 0$. 故 $\eta_n = 0$(a. s)($\forall\ n\in \mathrm{N}$). 由此即得唯一性.

定义 4.7.1 设$\zeta\in M^s(P,\mathscr{F}.)$,且非负. 若 $E\zeta_n\to 0$,则称ζ为势.

定理 4.7.2(Doob 分解) 若ζ为势,则$\exists$ 一个可料增序列 $\alpha = (\alpha_n, n\in N)(\mathscr{F}_{-1} = \mathscr{F}_0)$, $\ni \alpha_\infty \in L^1$,且

$$\zeta_n = E(\alpha_\infty/\mathscr{F}_n) - \alpha_n (\forall\ n\in \mathrm{N}).$$

这样的分解是唯一的.

证 设 $\Delta\zeta_n = \zeta_n - \zeta_{n-1}$, $\alpha_0 = 0$, $\alpha_n = -\sum_{k=1}^{n} E(\Delta\zeta_k/\mathscr{F}_{k-1})$($\forall\ n\geqslant 1$),则由$\zeta$的上鞅性知: $\alpha_n\uparrow$(a. s). 记 $\alpha_\infty = $(a. s)$-\lim_{n\to\infty}\alpha_n$,则按 Fatou 引理,有

$$E\alpha_\infty \leqslant \lim_{n\to\infty}E\alpha_n = \lim_{n\to\infty}(-\sum_{k=1}^{n}EE(\Delta\zeta_k/\mathscr{F}_{k-1}))$$
$$= \lim_{n\to\infty}(E\zeta_0 - E\zeta_n) = E\zeta_0 < \infty.$$

故 $\alpha_\infty\in L^1$. 显然, α_n 是$\mathscr{F}_{n-1}$-可测的. 因此, α为可料增序列.

现在，令 $\eta_n=\zeta_n+\alpha_n$，则 $\eta\in M(P,\mathscr{F}.)$. 事实上，对 $\forall\ n\in\mathrm{N}$，有

$$E(\eta_{n+1}/\mathscr{F}_n)=E(\zeta_{n+1}-\sum_{k=1}^{n+1}E(\Delta\zeta_k/\mathscr{F}_{k-1})/\mathscr{F}_n)$$
$$=\eta_n+E(\Delta\zeta_{n+1}-E(\Delta\zeta_{n+1}/\mathscr{F}_n)/\mathscr{F}_n)=\eta_n.$$

显然，ζ 和 α 都是一致可积的. 故 η 一致可积. 按定理 4.5.3，$\exists\ \eta_\infty\in L^1$，且 η_∞ 是 $\mathscr{F}_\infty$-可测的，使得

$$\eta_n=E(\eta_\infty/\mathscr{F}_n)，且\ \eta_n\xrightarrow{\text{a.s}}\eta_\infty.$$

$$\Rightarrow\eta_\infty=(\text{a.s})-\lim_{n\to\infty}\eta_n=(\text{a.s})-\lim_{n\to\infty}\zeta_n+(\text{a.s})-\lim_{n\to\infty}\alpha_n=\alpha_\infty.$$

（按定理 4.5.2，$(\text{a.s})-\lim\limits_{n\to\infty}\zeta_n=0$）. 综上分析，得

$$\zeta_n=E(\alpha_\infty/\mathscr{F}_n)-\alpha_n(\forall\ n\in\mathrm{N}).$$

为证唯一性，设 α'，α'' 都满足定理中的要求. 令 $\bar{\alpha}=\alpha'-\alpha''$，则

$$\bar{\alpha}_n=E(\alpha''_\infty-\alpha'_\infty/\mathscr{F}_n)=E(-\bar{\alpha}_\infty/\mathscr{F}_n).$$

注意到 $\bar{\alpha}$ 是可料的. 于是

$$\bar{\alpha}_{n-1}=E(-\bar{\alpha}_\infty/\mathscr{F}_{n-1})=E(E(-\bar{\alpha}_\infty/\mathscr{F}_n)/\mathscr{F}_{n-1})$$
$$=E(\bar{\alpha}_n/\mathscr{F}_{n-1})=\bar{\alpha}_n.\ (\forall\ n\geqslant 1)$$
$$\Rightarrow\bar{\alpha}_n=\bar{\alpha}_0=0，即\ \alpha'=\alpha''(\text{a.s}).$$

定义 4.7.2 设 $\alpha\in A_s(\mathscr{F}.)$，$\alpha_0=0$，$\alpha_n\uparrow\alpha_\infty(\text{a.s})$. 若 $\alpha_\infty\in L^1$，则称 α 为可积增序列. 若还可料，则称 α 为可料可积增序列.

从势的 Doob 分解中可以看出：势由两个因素决定，其一为 σ-代数流 $(\mathscr{F}.)$；其二为可料可积增序列 α. 因此，讨论 $\alpha\in A_s(\mathscr{F}.)$ 的可料性就成为一个比较重要的问题.

定理 4.7.3 可积增过程 α 具有可料性的充要条件是：对任意有界鞅 $\eta\in M(P,\mathscr{F}.)$，有

$$E(\sum_{n=0}^{\infty}\eta_n\cdot(\alpha_{n+1}-\alpha_n))=E(\eta_\infty\cdot\alpha_\infty),$$

其中 $\qquad\eta_\infty=\overline{\lim_{n\to\infty}}\eta_n，\alpha_\infty=(\text{a.s})-\lim\limits_{n\to\infty}\alpha_n.$

证 按 Levy 定理 4.5.3，对任意有界鞅 η，有

$$\eta_\infty = (\text{a.s}) - \lim_{n\to\infty}\eta_n，且\ \eta_n = E(\eta_\infty/\mathscr{F}_n)\ (\forall\ n\in \mathrm{N}).$$

按 Lebesque 控制收敛定理，有

$$\lim_{n\to\infty}E(\eta_n\alpha_{n+1}) = E(\eta_\infty\cdot\alpha_\infty).$$

$$E(\sum_{n=0}^{\infty}\eta_n(\alpha_{n+1}-\alpha_n)) = \lim_{m\to\infty}E(\sum_{n=0}^{m}\eta_n(\alpha_{n+1}-\alpha_n))$$

$$= \lim_{n\to\infty}[E(\sum_{n=0}^{m-1}\alpha_{n+1}(\eta_n-\eta_{n+1})) + E(\eta_m\cdot\alpha_{m+1})]$$

$$= \gamma + E(\eta_\infty\cdot\alpha_\infty),$$

其中
$$\gamma = \lim_{m\to\infty}E(\sum_{n=0}^{m-1}\alpha_{n+1}(\eta_n-\eta_{n+1})).$$

于是，α 可料的充要条件是：$\gamma=0$.

若 α 可料，则由 η 的鞅性质，立得 $\gamma=0$.

反之，假如 $\gamma=0$. 对任意有界鞅 η 成立. 选如下 η：

$$\eta_k = \begin{cases}\nu - E(\nu/\mathscr{F}_{n-1}), & 当\ k\geqslant n\ 时,\\ 0, & 当\ k<n\ 时.\end{cases}$$

其中 ν 为 $\mathscr{F}_n$-可测的有界 r. v. 显然，η 为有界鞅. 按条件，对此 η 应有 $\gamma=0$. 于是，

$$E[(\alpha_n - E(\alpha_n/\mathscr{F}_{n-1}))(\nu - E(\nu/\mathscr{F}_{n-1}))] = 0$$

$$\Rightarrow E\nu(\alpha_n - E(\alpha_n/\mathscr{F}_{n-1})) = 0.$$

现在，选取 $\nu=\text{sign}(\alpha_n-E(\alpha_n/\mathscr{F}_{n-1}))$（符号算子），并代入上式，得

$$E|\alpha_n - E(\alpha_n/\mathscr{F}_{n-1})| = 0$$

$$\Rightarrow \alpha_n = E(\alpha_n/\mathscr{F}_{n-1})(\text{a.s}),即\ \alpha\ 可料.$$

推论 4.7.1 可积增序列 α 具有可料性的充要条件是，对$\forall$ 有界鞅 η，有

$$E(\sum_{n=0}^{\infty}\eta_n\cdot\Delta\alpha_{n+1}) = E(\sum_{n=0}^{\infty}\eta_{n+1}\cdot\Delta\alpha_{n+1}),$$

其中
$$\Delta\alpha_{n+1} = \alpha_{n+1} - \alpha_n.$$

证 按定理 4.7.3，此命题等价于证明：

$$E\left(\sum_{n=0}^{\infty}\eta_{n+1}\Delta\alpha_{n+1}\right)=E(\eta_{\infty}\cdot\alpha_{\infty}).$$

按 Lebesque 控制收敛定理，有

$$\begin{aligned}E\left(\sum_{n=0}^{\infty}\eta_{n+1}\Delta\alpha_{n+1}\right)&=\lim_{m\to\infty}E\left(\sum_{n=0}^{m}\eta_{n+1}\Delta\alpha_{n+1}\right)\\&=\lim_{m\to\infty}\sum_{n=0}^{m}[E(\eta_{n+1}\cdot\alpha_{n+1})-E(\eta_{n}\cdot\alpha_{n})]\\&=\lim_{m\to\infty}E(\eta_{m+1}\cdot\alpha_{n+1})=E(\eta_{\infty}\cdot\alpha_{\infty}).\end{aligned}$$

定义 4.7.3 称 Doob 分解中的可料可积增序列 α 为势 ζ 的相联过程(序列). α 与 ζ 为互联过程.

推论 4.7.2 若 $\xi\in M^{s}(P,\mathscr{F}.)$，且一致可积，则 ξ 可唯一地表示成如下形式：$\xi_n=\eta_n-\alpha_n(\forall\ n\in\mathrm{N})$，其中 η 为一致可积鞅，α 为可料可积增序列.

证 将 Riesz 分解与 Doob 分解结合起来即得所要的结论.

关于上鞅分解定理的最重要应用在于研究平方可积鞅的性质，而平方可积鞅在鞅论中占有相当重要的地位. 下一节专门讨论这类鞅.

§4.8 平方可积鞅和特征

本节所考虑的随机变量都是定义在完全概率空间$(\Omega,\mathscr{F},P)$上. 所引用的某些符号凡没有重申的均按上一节约定.

定义 4.8.1 若 $\xi\in M(P,\mathscr{F}.)$，且 L^2-有界，则称 ξ 为平方可积鞅. (L^2-有界$\Leftrightarrow\sup\limits_{n}E|\xi_n|^2<\infty$).

引理 4.8.1 若 ξ 为平方可积鞅，则

$$E(\sup_{n}|\xi_n|^2)\leqslant 4\sup_{n}E|\xi_n|^2<\infty.$$

证 按定理 4.1(ii)，对$\forall\ n\in\mathrm{N}$，有

$$E(\max_{0\leqslant k\leqslant n}|\xi_k|^2)\leqslant 4E|\xi_n|^2\leqslant 4\sup_{n}E|\xi_n|^2.$$

让 $n\uparrow\infty$，由上式即得所要的结论.

推论 4.8.1 若 ξ 为平方可积鞅，则 ξ 可唯一地表示成如下形式：

$$\xi_n^2 = \zeta_n + \alpha_n \quad (\forall\ n \in \mathrm{N}),$$

其中 ζ 为一致可积鞅，α 为可料可积增序列.

证 按引理 4.8.1，ξ^2 为一致可积的非负下鞅. 从而，$-\xi^2$ 为一致可积上鞅. 按推论 4.7.2，$-\xi_n^2=\eta_n-\alpha_n$. 令 $\zeta_n=-\eta_n$，则 $\xi_n^2=\zeta_n+\alpha_n(\forall\ n\in\mathrm{N})$.

定义 4.8.2 设 $\xi\in M(P,\mathscr{F}.)$，则称与 ξ^2 相联的可料可积增序列 α 为鞅 ξ 的特征. 记为 $\alpha_n=\langle\xi\rangle_n$ 或 $\langle\xi,\xi\rangle n(\forall\ n\in\mathrm{N})$.

推论 4.8.2 若 ξ 为平方可积鞅，则

$$\Delta\langle\xi\rangle_n = E((\Delta\xi_n)^2/\mathscr{F}_{n-1})(\mathrm{a.s})(\forall\ n \geqslant 1).$$

证 按推论 4.8.1，可得

$$\begin{aligned} E((\Delta\xi_n)^2/\mathscr{F}_{n-1}) &= E(\Delta\xi_n^2/\mathscr{F}_{n-1}) \\ &= E(\Delta\zeta_n + \Delta\alpha_n/\mathscr{F}_{n-1}) \\ &= E(\Delta\alpha_n/\mathscr{F}_{n-1}) = \Delta\alpha_n = \Delta\langle\xi\rangle_n. \end{aligned}$$

推论 4.8.3 若 ξ、η 均为平方可积鞅，则乘积 $\xi\cdot\eta=(\xi_n\cdot\eta_n, n\in\mathrm{N})$ 可唯一地表示成如下形式：

$$\xi\cdot\eta = \gamma_n + \bar{\alpha}_n, \bar{\alpha}_n = \bar{\alpha}'_n - \bar{\alpha}''_n(\forall\ n \in \mathrm{N})$$

其中 γ 为一致可积鞅，$\bar{\alpha}'$、$\bar{\alpha}''$ 为可料可积增序列.

证 事实上

$$\xi_n\cdot\eta_n = \frac{1}{2}[\xi_n^2 + \eta_n^2 - (\xi_n - \eta_n)^2].$$

注意，ξ-η 仍为平方可积鞅. 因此，按推论 4.8.1，有

$$\xi_n\cdot\eta_n = \frac{1}{2}(\zeta'_n + \zeta''_n - \zeta'''_n) + \frac{1}{2}(\alpha'_n + \alpha''_n - \alpha'''_n).$$

此式右边第一项为一致可积鞅，记为 γ_n. 第二项为可料序列. 记为 $\bar{\alpha}_n$. 而让 $\bar{\alpha}'_n=\frac{1}{2}(\alpha'_n+\alpha''_n)$，$\bar{\alpha}''_n=\frac{1}{2}\alpha'''_n$，则 $\bar{\alpha}_n=\bar{\alpha}'_n-\bar{\alpha}''_n$. 合起来，即得推论中的表示.

定义 4.8.3 称推论 4.8.3 中的 $\bar{\alpha}$ 为 ξ、η 的联合特征. 记为 $\bar{\alpha} = \langle \xi, \eta \rangle$. (显然,联合特征可以表成两个可料可积增序列之差,因此,$\bar{\alpha}$ 为具有有限变差的可料序列).

推论 4.8.4 设 ξ、η 为平方可积鞅,$\langle \xi, \eta \rangle$ 为它们的联合特征. 则

$$\Delta\langle \xi, \eta \rangle_n = E(\Delta\xi_n \cdot \Delta\eta_n / \mathscr{F}_{n-1})(\mathrm{a.s})(\forall\, n \in \mathrm{N}).$$

证 利用推论 4.8.3 及 ξ、η 的鞅性质即可得此结论.

为了进一步研究平方可积鞅的性质,下面讨论半鞅的收敛集特点及结构. 首先定义下列符号:

$$\{\xi_n \to\} \triangleq \{-\infty < \lim_{n\to\infty} \xi_n < +\infty\},$$

$$\mathscr{U}^+ = \{\xi : \xi \in A_s(\mathscr{F}.), E((\Delta\xi_{\tau_a})^+ 1_{(\tau_a<\infty)}) < \infty (\forall\, a > 0)\}$$

其中 $\xi_{-1}=0, \mathscr{F}_{-1}=\mathscr{F}_0, \tau_a=\inf\{n\in\mathrm{N}: \xi_n > a\}, \inf\{\varnothing\}=\infty$.

注 1 "$\xi \in \mathscr{U}^+$"$\Leftrightarrow$"$E(\xi^+_{\tau_a} 1_{(\tau_a<\infty)}) < \infty (\forall\, a > 0)$".

事实上,

$$\begin{aligned} E(\xi^+_{\tau_a} 1_{(\tau_a<\infty)}) &\leqslant E[((\Delta\xi_{\tau_a})^+ + \xi^+_{\tau_a-1}) 1_{(\tau_a<\infty)}] \\ &\leqslant a + E((\Delta\xi_{\tau_a})^+ 1_{(\tau_a<\infty)}) \\ &\leqslant 2a + E(\xi^+_{\tau_a} 1_{(\tau_a<\infty)}). \end{aligned}$$

由此即知:等价关系成立.

定理 4.8.1 若 $\xi \in \mathscr{U}^+$,且为下鞅,即 $\xi \in M_s(P, \mathscr{F}.) \cap \mathscr{U}^+$,则

$$\{\xi_n \to\} = \{\sup_n \xi_n < \infty\} (\mathrm{mod} P).$$

证 显然,$\{\xi_n \to\} \subset \{\sup_n \xi_n < \infty\}$. 下证反向包含关系.

令 $\xi^{\tau_a} = (\xi_{n\wedge\tau_a}, n\in\mathrm{N}), (\mathscr{F}^{\tau_a}.) = (\mathscr{F}_{n\wedge\tau_a}, n\in\mathrm{N})$,则

ξ^{τ_a}、$(\xi^{\tau_a})^+ \in M_s(P, \mathscr{F}^{\tau_a}.) \cap \mathscr{U}^+$.

按注 1 及 Fatou 引理,有

$$\begin{aligned} E\xi^+_{n\wedge\tau_a} &= E(\xi^+_{n\wedge\tau_a}(1_{(\tau_a<\infty)} + 1_{(\tau_a=\infty)})) \\ &\leqslant a + E(\xi^+_{\tau_a} 1_{(\tau_a<\infty)}) < \infty. \end{aligned}$$

这表明：$(\xi^{\tau_a})^+$为L^1-有界非负下鞅. 因此，按 Doob 鞅收敛定理，$\xi_{n\wedge\tau_a}\xrightarrow{\text{a.s.}}\xi_{\tau_a}\in L^1$.

$$\Rightarrow(\tau_a=\infty)\subseteq\{\xi_n\rightarrow\}(\text{mod}P)(\forall\ a>0).$$

$$\Rightarrow\{\sup_n\xi_n<\infty\}=\bigcup_{a>0}(\tau_a=\infty)\subseteq\{\xi_n\rightarrow\}(\text{mod}P).$$

推论 4.8.5 设$\xi\in M(P,\mathscr{F}.)$，$\xi_{-1}=0$，$\mathscr{F}_{-1}=\mathscr{F}_0$，且$E(\sup_n|\Delta\xi_n|)<\infty$，则

$$\{\xi_n\rightarrow\}\cup\{\underline{\lim}\xi_n=-\infty,\overline{\lim}\xi_n=+\infty\}=\Omega\ (\text{mod}P).$$

证 按假设，易证：$\pm\xi\in\mathscr{U}^+$. 按定理 4.8.1，有

$$\{\overline{\lim}\xi_n<\infty\}=\{\sup_n\xi_n<\infty\}=\{\xi_n\rightarrow\}(\text{mod}P).$$

$$\{\underline{\lim}\xi_n>-\infty\}=\{\sup_n(-\xi_n)<\infty\}=\{\xi_n\rightarrow\}(\text{mod}P).$$

$$\Rightarrow\{\xi_n\rightarrow\}=\{\overline{\lim}\xi_n<\infty\}\cup\{\underline{\lim}\xi_n>-\infty\}\ (\text{mod}P).$$

由此，立即可得所要的结果.

定理 4.8.2（Riesz-Doob 分解定理的推广）. 设$\xi\in M^s(P,\mathscr{F}.)$，则$\xi$可唯一地表示成如下形式：

$$\xi_n=\eta_n+\alpha_n(\forall\ n\in\mathrm{N}),$$

其中$\eta\in M(P,\mathscr{F}.)$，$\alpha$为可料增序列.

证 设$\sigma_n=n$，则$\sigma_n\uparrow\infty$. 按 Doob 停时代换定理，$\xi^{\sigma_n}\in M^s(P,\mathscr{F}^{\sigma_n}.)$，且$L^1$-有界. 按 Riesz-Doob 分解定理，$\xi^{\sigma_n}$可唯一地表示为

$$\xi_{k\wedge\sigma_n}=\eta_{k\wedge\sigma_n}+\alpha_{k\wedge\sigma_n},$$

其中η^{σ_n}为一致可积鞅，α^{σ_n}为可料可积增序列. 按唯一性，当$k\leqslant n$时，$\eta_{k\wedge\sigma_n}$、$\alpha_{k\wedge\sigma_n}$关于n不变. 令

$$\eta_k=\lim_{n\to\infty}\eta_{k\wedge\sigma_n},\alpha_k=\lim_{n\to\infty}\alpha_{k\wedge\sigma_n},$$

则$\eta\in M(P,\mathscr{F}.)$，$\alpha$为可料增序列，且$\xi_k=\eta_k+\alpha_k(\forall\ k\in\mathrm{N})$.

分解的唯一性是显然的.

注 2 在 R.-D. 定理的分解式中，α是可积的. 推广的结果则不要求α是可积的. 此结果还可以进一步推广到局部上鞅的情形.

定理 4.8.3 设$\xi\in M_s(P,\mathscr{F}.)$，且 R.-D. 分解为

$$\xi_n = \eta_n + \alpha_n (\forall\, n \in \mathrm{N}),$$

其中$\eta \in M(P, \mathscr{F}.)$，$\alpha$为可料增序列，且$\alpha_n \in L'$ $(\forall\, n \in \mathrm{N})$.

(i) 若ξ非负，则$\{\alpha_\infty < \infty\} \subseteq \{\xi_n \to\} \subseteq \{\sup_n \xi_n < \infty\}$ (mod$P$).

(ii) 若$\xi \in \mathscr{U}^+$，则$\{\xi_n \to\} = \{\sup_n \xi_n < \infty\} \subseteq \{\alpha_\infty < \infty\}$ (mod$P$).

(iii) 若$\xi \in \mathscr{U}^+$，且非负，则$\{\alpha_\infty < \infty\} = \{\xi_n \to\} = \{\sup_n \xi_n < \infty\}$ (mod$P$).

证　若(i)、(ii)成立，则(iii)显然成立.

先证(i). 显然，$\{\xi_n \to\} \subseteq \{\overline{\lim}\xi_n < \infty\} = \{\sup_n \xi_n < \infty\}$. 为证(i)中的另一包含关系，定义

$$\sigma_a = \inf\{n \in \mathrm{N}: \alpha_{n+1} > a\}, \inf\{\varnothing\} = \infty \ (a > 0).$$

由α的可料性知：σ_a为停时，且$\alpha_{\sigma_a} \leqslant a$. 按Doob停止代换定理，有

$$E\xi_{n\wedge\sigma_a} = E\eta_{n\wedge\sigma_a} + E\alpha_{n\wedge\sigma_a} \leqslant E\xi_0 + a < \infty.$$

这表明：ξ^{σ_a}为L^1-有界下鞅. 按Doob收敛定理，ξ_{σ_a}为r.v. 从而$(\sigma_a = \infty) \subseteq \{\xi_n \to\}$ (mod$P$) $(\forall\, a > 0)$. 显然，$(\sigma_a = \infty) = \{\alpha_\infty \leqslant a\}$

$$\Rightarrow \{\alpha_\infty < \infty\} = U_{a>0}\{\alpha_\infty \leqslant a\} = U_{a>0}(\sigma_a = \infty)$$
$$\subseteq \{\xi_n \to\} (\mathrm{mod}P).$$

现在证(ii). 按定理4.8.1，有

$$\{\xi_n \to\} = \{\sup_n \xi_n < \infty\} (\mathrm{mod}P).$$

因此，为证(ii)，仅需证明其中的第二包含关系. 按Doob停时代换定理，有

$$E\alpha_{n\wedge\tau_a} = E\xi_{n\wedge\tau_a} - E\eta_{n\wedge\tau_a}$$
$$\leqslant -E\xi_0 + a + E((\Delta\xi_{\tau_a})^+ 1_{(\tau_a < \infty)}) < \infty.$$

由此可知：α^{τ_a}为L'-有界增序列. 按单调收敛定理，

$$E\alpha_{\tau_a} = \lim_{n\to\infty} E\alpha_{n\wedge\tau_a} < \infty, 即\ \alpha_{\tau_a} \in L' \ (\forall\, a > 0)$$
$$\Rightarrow (\tau_a = \infty) \subseteq \{\alpha_\infty < \infty\}.$$
$$\Rightarrow \{\sup_n \xi_n < \infty\} = U_{a>0}(\tau_a = \infty) \subseteq \{\alpha_\infty < \infty\} (\mathrm{mod}P).$$

推论 4.8.6 设$\xi\in A_s(\mathscr{F}.)$，且非负. $\xi_n\in L'$ ($\forall\ n\in\mathrm{N}$)，则

$$\{\sum_{n=1}^{\infty}E(\xi_n/\mathscr{F}_{n-1})<\infty\}\subseteq\{\gamma_n\rightarrow\}(\mathrm{mod}P),\gamma_n=\sum_{k=0}^{n}\xi_k.$$

若$E(\sup\limits_n\xi_n)<\infty$，则

$$\{\sum_{n=1}^{\infty}E(\xi_n/\mathscr{F}_{n-1})<\infty\}=\{\gamma_n\rightarrow\}(\mathrm{mod}P).$$

证 令$\alpha_0=0,\alpha_n=\sum\limits_{k=1}^{n}E(\xi_k/\mathscr{F}_{k-1}),\eta_n=\gamma_n-\alpha_n(\forall\ n\in\mathrm{N})$，则$\eta\in M(P,\mathscr{F}.)$，$\alpha$为可料增序列，且$\alpha_n\in L'$ ($\forall\ n\in\mathrm{N}$). 按定理4.8.3(i)，本命题第一个结论成立. 注意，γ为非负下鞅，且$\gamma\in\mathscr{U}^+$当$E(\sup\limits_n\xi_n)<\infty$时. 故按定理4.8.3(iii)，本命题的第二个结论成立.

推论 4.8.7(Borel-Cantelli-Levy 引理) 设$A_0=\varnothing,A_n\in\mathscr{F}_n$ $(n\geqslant1)$，则

$$\{\sum_{n=1}^{\infty}P(A_n/\mathscr{F}_{n-1})<\infty\}=\{\sum_{n=1}^{\infty}1_{A_n}<\infty\}(\mathrm{mod}P).$$

证 这只要令$\xi_n=1_{A_n}$，则由推论4.8.6，立得结论.

定理 4.8.4 设$\xi\in M(P,\mathscr{F}.),\xi_n\in L^2(\forall\ n\in\mathrm{N})$，则

$$\{\langle\xi\rangle_\infty<\infty\}\subseteq\{\xi_n\rightarrow\}(\mathrm{mod}P).$$

若 $E(\sup\limits_n|\Delta\xi_n|^2)<\infty$，则$\{\langle\xi\rangle_\infty<\infty\}=\{\xi_n\rightarrow\}(\mathrm{mod}P)$.

证 按假设，$\xi^2\in M_s(P,\mathscr{F}.)$，$(\xi+1)^2\in M_s(P,\mathscr{F}.)$. 它们的R.-D.分解为

$$\xi_n^2=\eta'_n+\alpha'_n,(\xi_n+1)^2=\eta''_n+\alpha''_n(\forall\ n\in\mathrm{N}),$$

其中

$$\alpha'_n=\sum_{k=1}^{n}E(\Delta\xi_k^2/\mathscr{F}_{k-1}),\alpha''_n=\sum_{k=1}^{n}E(\Delta(\xi_k+1)^2/\mathscr{F}_{k-1}).$$

显然，$\alpha'_n=\alpha''_n$，即$\langle\xi\rangle_n=\alpha'_n=\alpha''_n$. 按定理4.8.3(i)，

$$\begin{aligned}\{\langle\xi\rangle_\infty<\infty\}&=\{\alpha'_\infty<\infty\}\cap\{\alpha''_n<\infty\}\\&\subseteq\{\xi_n^2\rightarrow\}\cap\{(\xi_n+1)^2\rightarrow\}=\{\xi_n\rightarrow\}(\mathrm{mod}P).\end{aligned}$$

此即第一结论. 为证第二结论，按定理4.8.3(iii)，仅需证明：$\xi^2\in$

$\mathscr{U}^+$. 从而，$(\xi+1)^2\in\mathscr{U}^+$. 为此，定义

$$\tau_a=\inf\{n\in \mathrm{N}:\xi_n^2>a\},\inf\{\varnothing\}=\infty,A_a=(\tau_a<\infty).$$

显然，对$\forall\ a>0$，有

$$|\Delta\xi_{\tau_a}^2|\cdot 1_{A_a}\leqslant[(\Delta\xi_{\tau_a})^2+2\sqrt{a}\,|\Delta\xi_{\tau_a}|]\cdot 1_{A_a}.$$

$$\begin{aligned}E(|\Delta\xi_{\tau_a}^2|\cdot 1_{A_a})&\leqslant E(|\Delta\xi_{\tau_a}|^2\cdot 1_{A_a})\\&\quad+2\sqrt{a}\,(E(|\Delta\xi_{\tau_a}|^2\cdot 1_{A_a}))^{\frac{1}{2}}\\&\leqslant E(\sup_n|\Delta\xi_n|^2)+2\sqrt{a}\,(E\sup_n|\Delta\xi_n|^2)^{\frac{1}{2}}<\infty\\&\Rightarrow\xi^2\in\mathscr{U}^+.\end{aligned}$$

定理 4.8.5 设$\xi\in M(P,\mathscr{F}.)$，$\xi^2\in M_s(P,\mathscr{F}.)$.

(i)若α为可料增序列，且满足下列要求：

$$\alpha_1\geqslant 1(\mathrm{a.s}),\alpha_n\uparrow\infty(\mathrm{a.s}),\sum_{i=1}^{\infty}\frac{\Delta\langle\xi\rangle_i}{\alpha_i^2}<\infty(\mathrm{a.s}),$$

则 $\frac{1}{\alpha_n}\xi_n\xrightarrow{\mathrm{a.s}}0$ 当 $n\to\infty$时.

(ii)若$\langle\xi\rangle_n\uparrow\infty(\mathrm{a.s})$，则$\frac{1}{\langle\xi\rangle_n}\xi_n\xrightarrow{\mathrm{a.s}}0$ 当 $n\to\infty$时.

证 令 $\eta_0=0,\eta_n=\sum_{k=1}^{n}\frac{\Delta\xi_k}{\alpha_k}(n\geqslant 1)$，则 $\eta\in M(P,\mathscr{F}.)$，$\eta^2\in M_s(P,\mathscr{F}.)$. 且$\langle\eta\rangle_n=\sum_{k=1}^{n}E(|\Delta\eta_k|^2/\mathscr{F}_{k-1})=\sum_{k=1}^{n}\alpha_k^{-2}\Delta\langle\xi\rangle_k$. 按(i)中的假设，$\{\langle\eta\rangle_\infty<\infty\}=\Omega(\mathrm{mod}P)$. 按定理 4.8.4，有$\{\langle\eta\rangle_\infty<\infty\}\subseteq\{\eta_n\to\}(\mathrm{mod}P)$. 故$\{\eta_n\to\}=\Omega(\mathrm{mod}P)$.

又
$$\frac{1}{\alpha_n}\xi_n=\frac{1}{\alpha_n}\sum_{k=1}^{n}\alpha_k\Delta\eta_k=\eta_n+\frac{1}{\alpha_n}\sum_{k=1}^{n-1}\eta_k\Delta\alpha_{k+1}.$$

按 Kronecker 引理，有$\frac{1}{\alpha_n}\xi_n\xrightarrow{\mathrm{a.s}}0$.

现在，考虑(ii). 令 $A_n=\{\langle\xi\rangle_n\geqslant 1\}\in\mathscr{F}_{n-1}$，则 $A_n\uparrow\Omega(\mathrm{mod}P)$. 令 $\alpha_n=\langle\xi\rangle_n 1_{A_n}+n1_{A_n^c}$，若 α 满足(i)中的要求，则$\frac{1}{\alpha_n}\xi_n\xrightarrow{\mathrm{a.s}}0$. 于是，对$\forall\ m\in N$，有

$$\frac{1}{\langle\xi\rangle_n}\xi_n 1_{A_m}=\frac{1}{\alpha_n}\xi_n 1_{A_m}\xrightarrow{\text{a.s}}0 \text{ 当 } n\to\infty \text{ 时}.$$

故$\frac{1}{\langle\xi\rangle_n}\xi_n\xrightarrow{\text{a.s}}0$ 于 Am 上($\forall\ m\in\mathrm{N}$). 由此即得结论(ii). 余下的问题就是验证 α 满足(i)中的要求. 显然, α 可料, $\alpha_1\geqslant 1$(a. s), $\alpha_n\uparrow\infty$(a. s). 关于(i)中级数的收敛性判断如下.

$$\sum_{k=1}^{\infty}\frac{\Delta\langle\xi\rangle_k}{\alpha_k^2}\leqslant\sum_{k=1}^{\infty}\frac{\Delta\langle\xi\rangle_k}{\langle\xi\rangle_k^2}1_{A_k}+\sum_{k=1}^{\infty}\frac{1}{k^2}1_{A_k^c}.$$

此式右边第二个级数收敛. 假如 $\omega\in A_m$, 则

$$\sum_{k=1}^{\infty}\frac{\Delta\langle\xi\rangle_k}{\langle\xi\rangle_k^2}1_{A_k}=\sum_{k=m}^{\infty}\frac{\Delta\langle\xi\rangle_k}{\langle\xi\rangle_k^2}+\sum_{k=1}^{m-1}\frac{\Delta\langle\xi\rangle_k}{\langle\xi\rangle_k^2}1_{A_k}.$$

此式右边级数的收敛性等价于以下初等命题:

$$\text{“}\sum_{n=1}^{\infty}\frac{a_n}{b_n^2}<\infty, \text{其中 } a_n>0, b_n=\sum_{k=1}^{n}a_k\text{”}$$

由 $m\in N$ 的任意性即知: $\sum_{k=1}^{\infty}\frac{\Delta\langle\xi\rangle_k}{\langle\xi\rangle_k^2}1_{A_k}<\infty$(a. s).

定理 4.8.6 设 $\xi\in M_s(P,\mathscr{F}.)$, R.-D. 分解为 $\xi=\eta+\alpha$($\eta\in M(P,\mathscr{F}.)$, α 为可料增序列). 若 $|\Delta\xi_n|\leqslant c$(a. s), 则

(i) $\{\langle\eta\rangle_\infty+\alpha_\infty<\infty\}=\{\xi_n\to\}$ (mod$P$);

(ii) $\{\sum_{n=1}^{\infty}E(|\Delta\xi_n|^2+\Delta\xi_n/\mathscr{F}_{n-1})<\infty\}=\{\xi_n\to\}$ (mod$P$).

证 在 ξ 的 R.-D. 分解中, 有 $\alpha_0=0$, $\eta_0=\xi_0$, $\alpha_n=\sum_{k=1}^{n}E(\Delta\xi_k/\mathscr{F}_{n-1}$, $\eta_n=\xi_n-\alpha_n=\sum_{k=1}^{n}(\Delta\xi_k-E(\Delta\xi_k/\mathscr{F}_{k-1}))$ ($n\geqslant 1$). 按条件, $\eta\in M(P,\mathscr{F}.)$, $\eta^2\in M_s(P,\mathscr{F}.)$, $E(\sup_n|\Delta\xi_n|)\leqslant c<\infty$. $E(\sup_n|\Delta\eta_n|^2)\leqslant 4c^2<\infty$. 于是, $\xi\in\mathscr{U}^+$. 按定理 4.8.4, $\{\langle\eta\rangle_\infty<\infty\}=\{\eta_n\to\}$(mod$P$). 按定理 4.8.3(ii), $\{\xi_n\to\}\subseteq\{\alpha_\infty<\infty\}$(mod$P$). 按 R.-D. 分解, $\{\xi_n\to\}\supseteq\{\eta_n\to\}\cap\{\alpha_\infty<\infty\}$. 综合上述分析, 得

$$\{\xi_n \to\} = \{\xi_n \to\} \cap \{\alpha_\infty < \infty\} = \{\xi_n \to\} \cap \{\alpha_\infty < \infty\} \cap \{\eta_n \to\}.$$
$$= \{\alpha_\infty < \infty\} \cap \{\eta_n \to\} = \{\langle\eta\rangle_\infty + \alpha_\infty < \infty\} \quad (\mathrm{mod}P).$$

故结论(i)成立. 下证结论(ii).

注意 $$\langle\eta\rangle_n = \sum_{k=1}^n E(|\Delta\eta_k|^2/\mathscr{F}_{k-1})$$
$$= \sum_{k=1}^n E(|\Delta\xi_k|^2 - |E(\Delta\xi_k/\mathscr{F}_{k-1})|^2/\mathscr{F}_{k-1}).$$
$$\{\alpha_\infty < \infty\} = \{\alpha_\infty < \infty\} \cap \{\sum_{k=1}^\infty |E(\Delta\xi_k/\mathscr{F}_{k-1})|^2 < \infty\}$$
$$\Rightarrow \{\alpha_\infty + \langle\eta\rangle_\infty < \infty\} = \{\alpha_\infty < \infty\} \cap \{\langle\eta\rangle_\infty < \infty\}$$
$$= \{\sum_{k=1}^\infty E(|\Delta\xi_k|^2/\mathscr{F}_{k-1}) < \infty\} \cap \{\alpha_\infty < \infty\}.$$
$$= \{\sum_{k=1}^\infty E(\Delta\xi_k + |\Delta\xi_k|^2/\mathscr{F}_{k-1}) < \infty\} (\mathrm{mod}P).$$

定理 4.8.7(Kolmogorov 三级数定理的推广) 设 $\xi \in A_s(\mathscr{F}.)$, c 为任意正常数,则级数 $\sum_k \xi_k$ 在集 A_c 上收敛,

$$A_c = \{\sum_n P(|\xi_n| \geqslant c/\mathscr{F}_{n-1}) < \infty\} \cap \{\sum_n E(\xi_n^c/\mathscr{F}_{n-1}) < \infty\}$$
$$\cap \{\sum_n V(\xi_n^c/\mathscr{F}_{n-1}) < \infty\},$$

其中

$$\xi_n^c = \xi_n 1_{\{|\xi_n| \leqslant c\}}, V(\mathscr{X}/\mathscr{F}_{n-1}) = E((X - E(X/\mathscr{F}_{n-1}))^2/\mathscr{F}_{n-1}).$$

证 令 $\gamma_n = \sum_{k=1}^n \xi_k$. 按推论 4.8.7,有

$$\{\sum_n P(|\xi_n| \geqslant c/\mathscr{F}_{n-1}) < \infty\} = \{\sum_n 1_{\{|\xi_n| \geqslant c\}} < \infty\}.$$
$$\Rightarrow A_c \cap \{\gamma_n \to\} = A_c \cap \{\sum_{k=0}^n \xi_k^c \to\} = A_c \cap \{\nu_n \to\},$$

其中 $\nu_n = \sum_{k=0}^n \eta_k$, $\eta_k = \xi_k^c - E(\xi_k^c/\mathscr{F}_{k-1})$, $\eta_0 = 0$, $\mathscr{F}_{-1} = \mathscr{F}_0$. 显然, $\nu \in M(P, \mathscr{F}.)$, $|\Delta\nu_n| = |\eta_n| \leqslant 2c$. 按定理 4.8.4,有

$$\{\langle\nu\rangle_\infty < \infty\} = \{\nu_n \to\}(\mathrm{mod}P).$$

另外，
$$\Delta\langle\nu\rangle_n = E(\eta_n^2/\mathscr{F}_{n-1}) = V(\xi_n^c/\mathscr{F}_{n-1}),$$

$$\langle\nu\rangle_\infty = \sum_{n=1}^{\infty} V(\xi_n^c/\mathscr{F}_{n-1}).$$

$\Rightarrow A_c \cap \{\gamma_n \to\} = A_c \cap \{\nu_n \to\} = A_c \cap \{\langle\nu\rangle_\infty < \infty\} = A_c.$

$\Rightarrow A_c \subseteq \{\gamma_n \to\}.(\mathrm{mod}P).$

定理 4.8.8(中心极限定理)　设 $\xi \in M(P,\mathscr{F}.)$，且$(\Delta\xi_n, n \geq 1)$有界. 若$\langle\xi\rangle_\infty = \infty(\mathrm{a.s})$，则

$$\frac{1}{\sqrt{t}}\xi_{\nu_t} \xrightarrow[t\to\infty]{d.} \tilde{N}(\tilde{N}\text{ 表正态变量})$$

其中 $\nu_t = \inf\{n \geq 1 : \langle\xi\rangle_n \geq t\}, \inf\{\varnothing\} = \infty.$

证　按条件$\langle\xi\rangle_\infty = \infty(\mathrm{a.s})$，$\nu_t$ 必为有限停时，$(\forall\ t \geq 0)$. 令 $S_n^2 = \langle\xi\rangle_n, \sigma_k^2 = \Delta\langle\xi\rangle_k = E(|\Delta\xi_k|^2/\mathscr{F}_{k-1})$，则由$(\Delta\xi_n, n \geq 1)$有界可知：$\exists$ r.v. C_t 取值于$(0,1]$，使得

$$S_{\nu_t-1}^2 + c_t^2\sigma_{\nu_t}^2 = t.$$

(注意：$S_{\nu_t-1}^2 < t, S_{\nu_t}^2 \geq t \Rightarrow 0 < c_t < 1$). 于是，

$$c_t^2 = \frac{t - S_{\nu_t-1}^2}{\sigma_{\nu_t}^2}, \text{且 } c_t^2 1_{(\nu_t=n)} = \frac{t - S_{n-1}^2}{\sigma_n^2} 1_{(\nu_t=n)}(\mathscr{F}_{n-1}\text{-可测}).$$

令 $Y_n = \Delta\xi_n (n \geq 1)$，则由 Y 有界可知

$$\frac{1}{\sqrt{t}}(1 - c_t)Y_{\nu_t} \xrightarrow[t\to\infty]{\mathrm{a.s}} 0.$$

于是，"$\frac{1}{\sqrt{t}}\xi_{\nu_t} \xrightarrow{d.} \tilde{N}$"等价于"$\frac{1}{\sqrt{t}}(c_t Y_{\nu_t} + \sum_{k=1}^{\nu_t-1} Y_k) \xrightarrow{d.} \tilde{N}$".

令 $Z_n = Y_n \cdot 1_{(\nu_t > n)} + c_t Y_n \cdot 1_{(\nu_t = n)}$，则 Z 具有下列性质：

(Z.1) $$\sum_{k=1}^{\infty} Z_k 1_{(\nu_t \geq k)} = \sum_{k=1}^{\infty} Z_k \sum_{n=k}^{\infty} 1_{(\nu_t=n)} = \sum_{n=1}^{\infty}\sum_{k=1}^{n} Z_k 1_{(\nu_t=n)}$$

$$= \sum_{n=1}^{\infty} 1_{(\nu_t=n)} \cdot (Z_n + \sum_{k=1}^{n-1} Z_k) = c_t Y_{\nu_t} + \sum_{k=1}^{\nu_t-1} Y_k.$$

(Z. 2) “$\frac{1}{\sqrt{t}}\sum_{k=1}^{\infty}Z_k\xrightarrow{d.}\tilde{N}$”⇔“$\frac{1}{\sqrt{t}}(c_tY_{\nu_t}+\sum_{k=1}^{\nu_t-1}Y_k)\xrightarrow{d.}\tilde{N}$”.

(Z. 3) $E(Z_k|\mathscr{F}_{k-1})=E(Y_k\cdot 1_{(\nu_t>k)})+c_tY_k1_{(\nu_t=k)}/\mathscr{F}_{k-1})=0.$

(Z. 4) $\tau_k^2\triangleq E(Z_k^2/\mathscr{F}_{k-1})=\sigma_k^2 1_{(\nu_t>k)}+c_t^2\sigma_{\nu_t}^2 1_{(\nu_t=k)}.$

$$\sum_{k=1}^{\infty}\tau_k^2=\sum_{k=1}^{\infty}c_t^2\sigma_k^2 1_{(\nu_t=k)}+\sum_{k=1}^{\infty}\sum_{n=k+1}^{\infty}\sigma_k^2 1_{(\nu_t=n)}$$

$$=c_t^2\sigma_{\nu_t}^2+\sum_{n=2}^{\infty}\sum_{k=1}^{n-1}\sigma_k^2 1_{(\nu_t=n)}$$

$$=c_t^2\sigma_{\nu_t}^2+\sum_{k=1}^{\nu_t-1}\sigma_k^2=s_{\nu_t-1}^2+c_t^2\sigma_{\nu_t}^2=t.$$

设 $N=(N_k,k\geqslant 1)$为具有标准正态分布的独立 r. v. 列，且独立于 $\mathscr{F}_\infty$，其中 $\mathscr{F}_\infty=\sigma(\bigcup_n\mathscr{F}_n)$.

(N. 1) N 自然独立于每个 $\mathscr{F}_n$.

(N. 2) 如果$(\Omega,\mathscr{F})$的容量不够，可按如下方式扩充. 让$(\Omega\times\Omega',\mathscr{F}\otimes\mathscr{F}',P\times P')$为新的概率空间，然后，扩充变量的定义域，使它们都在此空间上有定义. 令

$$T_n=\sum_{k=1}^{n}Z_k+\sum_{k=n+1}^{\infty}\tau_k\cdot N_k,T_0=\sum_{k=1}^{\infty}\tau_k\cdot N_k,$$

则 T 具有下列特点：

(T. 1) $E(\exp(is\frac{T_0}{\sqrt{t}})/\mathscr{F}_\infty)$

$$=E(\exp(i\frac{s}{\sqrt{t}}\sum_{k=1}^{\infty}\tau_kN_k)/\mathscr{F}_\infty)$$

$$=\exp\left(-\frac{s^2}{2t}\sum_{k=1}^{\infty}\tau_k^2\right)=\exp\left(-\frac{1}{2}s^2\right)\quad(\forall\ s\in R)$$

$$\Rightarrow\frac{1}{\sqrt{t}}T_0\xrightarrow[t\to\infty]{d.}\tilde{N}.$$

(T. 2) “$\frac{1}{\sqrt{t}}T_{\nu_t}\xrightarrow[t\to\infty]{d.}\tilde{N}$”等价于“$\frac{1}{\sqrt{t}}\sum_{k=1}^{\infty}Z_k\xrightarrow[t\to\infty]{d.}\tilde{N}$”. 事实上

$$E\left(\exp\left(i\frac{s}{\sqrt{t}}T_{\nu_t}\right)/\mathscr{F}_\infty\right)$$

$$=E\left(\exp\left(i\frac{s}{\sqrt{t}}\left(\sum_{k=1}^{\nu_t}Z_k\right.\right.\right.$$

$$\left.\left.\left.+\sum_{k=\nu_t+1}^{\infty}\tau_kN_k\right)\right)/\mathscr{F}_\infty\right)$$

$$=\exp\left(i\frac{s}{\sqrt{t}}\sum_{k=1}^{\nu_t}Z_k\right)\cdot\exp\left(-\frac{s^2}{2t}\sum_{k=\nu_t+1}^{\infty}\tau_k^2\right);$$

(T. 3) $$\text{“}\frac{1}{\sqrt{t}}T_0\xrightarrow[t\to\infty]{d.}\tilde{N}\text{”}\Leftrightarrow\text{“}\frac{1}{\sqrt{t}}T_{\nu_t}\xrightarrow[t\to\infty]{d.}\tilde{N}\text{”}$$

$$\Leftrightarrow\text{“}\left|E\left(\exp\left(i\frac{s}{\sqrt{t}}T_{\nu_t}\right)-\exp\left(i\frac{s}{\sqrt{t}}T_0\right)\right)\right|\xrightarrow[t\to\infty]{}0\text{”}.$$

现在,剩下的问题就是证明:(T. 3)中最后的极限关系成立.

$$\left|E\left(\exp\left(i\frac{s}{\sqrt{t}}T_{\nu_t}\right)-\exp\left(i\frac{s}{\sqrt{t}}T_0\right)\right)\right|$$

$$=\left|E\sum_{n=1}^{\nu_t}\left(\exp\left(i\frac{s}{\sqrt{t}}T_n\right)-\exp\left(i\frac{s}{\sqrt{t}}T_{n-1}\right)\right)\right|$$

$$=\left|E\sum_{n=1}^{\nu_t}E\left(\exp\left(i\frac{s}{\sqrt{t}}T_n\right)-\exp\left(i\frac{s}{\sqrt{t}}T_{n-1}\right)/\mathscr{F}_\infty\right)\right|$$

$$=\left|E\sum_{n=1}^{\nu_t}\left[\exp\left(i\frac{s}{\sqrt{t}}\sum_{k=1}^{n}Z_k\right)\cdot\exp\left(-\frac{s^2}{2t}\sum_{k=n+1}^{\infty}\tau_k^2\right)\right.\right.$$

$$\left.\left.-\exp\left(i\frac{s}{\sqrt{t}}\sum_{k=1}^{n-1}Z_k\right)\cdot\exp\left(-\frac{s^2}{2t}\sum_{k=n}^{\infty}\tau_k^2\right)\right]\right|$$

$$=\left|E\sum_{n=1}^{\nu_t}\left\{\exp\left(-\frac{s^2}{2t}\sum_{k=n+1}^{\infty}\tau_k^2\right)\cdot\exp\left(i\frac{s}{\sqrt{t}}\sum_{k=1}^{n-1}Z_k\right)\right.\right.$$

$$\left.\left.\cdot\left[\exp\left(i\frac{s}{\sqrt{t}}Z_n\right)-\exp\left(-\frac{s^2}{2t}\tau_n^2\right)\right]\right\}\right|$$

$$=\left|\sum_{m=1}^{\infty}\sum_{n=1}^{m}E\left\{1_{(\nu_t=m)}\exp\left(-\frac{s^2}{2t}\left(t-\sum_{k=1}^{n}\tau_k^2\right)\right)\cdot\exp\left(i\frac{s}{\sqrt{t}}\sum_{k=1}^{n-1}Z_k\right)\right.\right.$$

$$
\cdot \left[\exp\left(i \frac{s}{\sqrt{t}} Z_n \right) - \exp\left(-\frac{s^2}{2t} \tau_n^2 \right) \right] \Big\} \Big|
$$

$$
= \left| \sum_{n=1}^{\infty} E\left[1_{(\nu_t \geqslant n)} \exp\left(-\frac{s^2}{2t}\left(t - \sum_{k=1}^{n} \tau_k^2\right) + i \frac{s}{\sqrt{t}} \sum_{k=1}^{n-1} Z_k \right) \right.\right.
$$

$$
\left.\left. \cdot E\left(\exp\left(i \frac{s}{\sqrt{t}} Z_n \right) - \exp\left(-\frac{s^2}{2t} \tau_n^2 \right) / \mathscr{F}_{n-1} \right) \right] \right|
$$

$$
\leqslant \sum_{n=1}^{\infty} E \left| E(\exp\left(i \frac{s}{\sqrt{t}} Z_n \right) - \exp\left(-\frac{s^2}{2t} \tau_n^2 \right) / \mathscr{F}_{n-1} \right) \Big|
$$

$$
= \sum_{k=1}^{\infty} E \left| E\left(\exp\left(i \frac{s}{\sqrt{t}} Z_n \right) - E\left(\exp\left(i \frac{s}{\sqrt{t}} \tau_n N_n \right) / \mathscr{F}_{n-1} \right) / \mathscr{F}_{n-1} \right) \right|
$$

$$
= \sum_{n=1}^{\infty} E | E(\beta_{nt} / \mathscr{F}_{n-1}) |,
$$

其中
$$
\beta_{nt} = \exp\left(i \frac{s}{\sqrt{t}} Z_n \right) - \exp\left(i \frac{s}{\sqrt{t}} \tau_n N_n \right).
$$

令 $x = \frac{s}{\sqrt{t}} Z_n, y = \frac{s}{\sqrt{t}} \tau_n N_n$. 利用不等式：

$$
\left| e^{ix} - \left(1 + ix - \frac{1}{2} x^2 \right) \right| \leqslant \min\left(\frac{1}{6} |x|^3, |x|^2 \right),
$$

则可得

$$
|E(\beta_{nt} / \mathscr{F}_{n-1})| = |E(e^{ix} - (1 + ix - \frac{1}{2} x^2) + (1 + ix - \frac{1}{2} x^2)
$$

$$
- (1 + iy - \frac{1}{2} y^2) + (1 + iy - \frac{1}{2} y^2) - e^{iy} / \mathscr{F}_{n-1})|
$$

$$
= |E(e^{ix} - (1 + ix - \frac{1}{2} x^2) + (1 + iy - \frac{1}{2} y^2) - e^{iy} / \mathscr{F}_{n-1})|
$$

$$
\leqslant E\left(\frac{1}{6} \left(\frac{|s|}{\sqrt{t}} \right)^3 |z_n|^3 + \frac{1}{6} \left(\frac{|s|}{\sqrt{t}} \right)^3 |\tau_n N_n|^3 / \mathscr{F}_{n-1} \right)
$$

$$
\leqslant K |s|^3 t^{-\frac{3}{2}} (z_n^2 + K^1 \tau_n^2).
$$

注意，$EZ_n^2 = E\tau_n^2, t = \sum_{n=1}^{\infty} \tau_n^2$. 由上述估计式便可得

$$\sum_{n=1}^{\infty} E|E(\beta_{nt}/\mathscr{F}_{n-1})| \leqslant c|s|^3 \frac{1}{\sqrt{t}} \xrightarrow[t\to\infty]{} 0.$$

由此即得(T.3)中的最后极限关系.

注 3 证明初等不等式：

$$\left|e^{ix}-(1+ix-\frac{1}{2}x^2)\right| \leqslant \min\left(\frac{1}{6}|x|^3, |x|^2\right).$$

事实上，$e^{ix}=\sum_{k=0}^{n}\frac{1}{k!}(ix)^k+\frac{i^{n+1}}{n!}\int_0^x (x-s)^n e^{is}ds.$

$$\int_0^x (x-s)^{n-1}e^{is}ds = \frac{x^n}{n}+\frac{i}{n}\int_0^x (x-s)^n e^{is}ds$$

$$\Rightarrow e^{ix}=\sum_{k=0}^{n}\frac{(ix)^k}{k!}+\frac{i^n}{(n-1)!}\int_0^x (x-s)^{n-1}(e^{is}-1)ds$$

$$\Rightarrow \left|e^{ix}-\sum_{k=0}^{n}\frac{(ix)^k}{k!}\right| \leqslant \frac{2}{(n-1)!}\left|\int_0^x (x-s)^{n-1}ds\right| = \frac{2}{n!}|x|^n.$$

$$\left|e^{ix}-\sum_{k=0}^{n}\frac{(ix)^k}{k!}\right| \leqslant \frac{1}{n!}\int_0^x (x-s)^n ds = \frac{1}{(n+1)!}|x|^{n+1}$$

$$\Rightarrow \left|e^{ix}-\sum_{k=0}^{n}\frac{(ix)^k}{k!}\right| \leqslant \min\left(\frac{1}{(n+1)!}|x|^{n+1}, \frac{1}{n!}|x|^n\right).$$

参数估计问题. 作为应用例，考虑如下估计问题：设 $\xi=(\xi_n, n\in N)$ 为独立的 r.v. 列，$E\xi_n=0, V\xi_n=E\xi_n^2=v_n>0(\forall\, n\in \mathrm{N})$. 考虑具有噪声 ξ 的参数系统：

$$\eta_{n+1}=\theta\eta_n+\xi_{n+1}(n\in N).$$

假设系统的初态 η_0 独立于 ξ，$\theta\in R$ 为待估的未知参数. 现在要求在时刻 n 通过观测数据 $\eta_0, \eta_1, \cdots, \eta_n$ 来估计参数 $\theta\in R$.

设 $\mathscr{F}_n=\sigma(\eta_0,\eta_1,\cdots,\eta_n)=\sigma(\eta_0,\xi_1,\cdots,\xi_n)$，则任意 $\mathscr{F}_n$-可测的函数 $\tilde{\theta}_n$ 均可看作是 θ 在时刻 n 的一个估计，而最佳估计问题是：要求在时刻 n 找到一个估计 $\hat{\theta}_n$，使得

$$f(\hat{\theta}_n)=\min\{f(\tilde{\theta}_n):\tilde{\theta}_n \text{ 为 } \theta \text{ 在时刻 } n \text{ 的任意估计}\}$$

其中 $f(\tilde{\theta}_n)=\sum_{k=0}^{n-1}c_k E(\eta_{k+1}-\tilde{\eta}_{k+1})^2, \tilde{\eta}\theta_{k+1}=\tilde{\theta}_n\eta_k$，

$c=(c_k, k\in N)$ 为加权系数(非负).

为解此问题，作如下变分：

$$f(\hat{\theta}_n+\delta\theta_n)-f(\hat{\theta}_n)=\sum_{k=0}^{n-1}c_k\cdot[E((\theta-\hat{\theta}_n-\delta\theta_n)\eta_k+\xi_{k+1})^2$$

$$-E((\theta-\hat{\theta}_n)\eta_k+\xi_{k+1})^2]$$

$$=-\sum_{k=0}^{n-1}c_k\cdot E[(2(\theta-\hat{\theta}_n)\eta_k$$

$$+2\xi_{k+1})\eta_k\delta\theta_n]+O(\delta\theta_n)^2$$

$$\Rightarrow\sum_{k=0}^{n-1}c_k\cdot E((\theta-\hat{\theta}_n)\eta_k^2+\eta_k\xi_{k+1})=0$$

$$\Rightarrow\hat{\theta}_n=\theta+\frac{\gamma_n}{\alpha_n}\ (n\geqslant1).$$

其中 $\gamma_n=\sum_{k=0}^{n-1}c_k\cdot\eta_k\xi_{k+1}$，$\alpha_n=\sum_{k=0}^{n-1}c_k\cdot\eta_k^2(n\geqslant1)$.

规定 $\gamma_0=0=\alpha_0$，则 $\gamma\in M(P,\mathscr{F}.)$，$\alpha$ 为可料增序列，且

$$\langle\gamma\rangle_n=\sum_{k=0}^{n-1}c_k^2\cdot v_{k+1}\cdot\eta_k^2(n\geqslant1).$$

定义 4.8.4 若估计 $\hat{\theta}=(\hat{\theta}_n,n\in N)$ (a.s)收敛到真值 θ，则称 $\hat{\theta}$ 为强相容估计，即 $P(\hat{\theta}_n\to\theta)=1$.

命题 4.8.1 若加权系数 $c_k=v_{k+1}^{-1}$，且 $\langle\gamma\rangle_n\uparrow\infty$(a.s)，则 $\hat{\theta}_n\xrightarrow[n\to\infty]{}\theta$(a.s).

证 事实上，$\gamma_n=\sum_{k=0}^{n-1}v_{k+1}^{-1}\eta_k\cdot\xi_{k+1}$，$\gamma\in M(P,\mathscr{F}.)$，$\gamma^2\in M_s(P,\mathscr{F}.)$. 按定理 4.8.5，$\frac{1}{\langle\gamma\rangle_n}\gamma_n\xrightarrow[n\to\infty]{}0$ (a.s). 当 $c_k=v_{k+1}^{-1}$ 时，$\alpha_n=\langle\gamma\rangle_n$.

故 $\hat{\theta}_n=\theta+\frac{\gamma_n}{\alpha_n}\xrightarrow[n\to\infty]{}\theta$ (a.s).

命题 4.8.2 若 $\sup_n\left\{\frac{v_{n+1}}{v_n}\right\}<\infty$，且 $\sum_n E((v_n^{-1}\xi_n^2)\wedge1)=\infty$，则 $\hat{\theta}$ 是强相容的.

证　令$\zeta_n=\frac{1}{v_n}\xi_n^2$,则

$$\begin{aligned}\sum_n 1\wedge\zeta_n &\leqslant \sum_n \zeta_n = \sum_n \frac{(\eta_n-\theta\eta_{n-1})^2}{v_n}\\ &\leqslant 2\sum_n\left(\theta^2+\frac{v_{n+1}}{v_n}\right)\frac{1}{v_{n+1}}\eta_n^2\\ &\leqslant 2\left(\theta^2+\sup_n\left\{\frac{v_{n+1}}{v_n}\right\}\right)\sum_n\frac{1}{v_{n+1}}\eta_n^2\\ &=2\left(\theta^2+\sup_n\left\{\frac{v_{n+1}}{v_n}\right\}\right)\langle\gamma\rangle_\infty\\ &\Rightarrow\{\sum_n 1\wedge\zeta_n=\infty\}\subseteq\{\langle\gamma\rangle_\infty=\infty\}.\end{aligned}$$

按K.氏三级数定理,有$P(\sum_n 1\wedge\zeta_n=\infty)=1$.事实上,在定理4.8.7中取$c=2$,以$1\wedge\zeta_n$代替$\xi_n$.注意到独立性,便得$\sum_n 1\wedge\zeta_n$收敛于集$A_c=\{\sum_n 1\wedge\zeta_n<\infty\}\cap\{\sum_n E(1\wedge\zeta_n)<\infty\}\cap\{\sum_n V(1\wedge\zeta_n)<\infty\}$上.注意,$V(1\wedge\zeta_n)=E(1\wedge\zeta_n-E(1\wedge\zeta_n))^2$.则$q\in M(P,\mathscr{F}.)$,$q^2\in M_s(P,\mathscr{F}.)$,$E(\sup_n|\Delta q_n|^2)<\infty$,其中$q_n=\sum_{k=1}^n(1\wedge\zeta_k-E(1\wedge\zeta_k))$.按定理4.8.4,有$\{\langle q\rangle_\infty<\infty\}=\{q_n\rightarrow\}$.而$\langle q\rangle_\infty=\sum_{n=1}^\infty V(1\wedge\zeta_n)$.故

$$A_c=\{\sum_n 1\wedge\zeta_n<\infty\}\cap\{\sum_n E1\wedge\zeta_n<\infty\}.$$

因此,若$\sum_n E1\wedge\zeta_n<\infty$,则$A_c=\{\sum_n 1\wedge\zeta_n<\infty\}\ (c=2)$.
按假设,得$P(A_c)=0$.从而

$$\begin{aligned}&P(\sum_n 1\wedge\zeta_n=\infty)=1\\ &\Rightarrow\langle\gamma\rangle_\infty=\infty\ (\text{a.s}).\end{aligned}$$

按命题4.8.1,便得所要的结论.

如何保证命题4.8.2中的条件成立的问题将在下一节讨论,并将假定ξ是Gauss的.

§ 4.9 鞅与分布的绝对连续性

设$(\Omega,\mathscr{F})$为可测空间. $(\mathscr{F}_.)=(\mathscr{F}_n,n\in\mathrm{N})$为$\mathscr{F}$中的子$\sigma$-代数流. $N=\{0,1,2,\cdots\}$. P、$\tilde{P}$为$(\Omega,\mathscr{F})$上的概率测度. $\mathscr{F}=\sigma(\bigcup_n\mathscr{F}_n)$. $P_n=P|_{\mathscr{F}_n}$和$\tilde{P}_n=\tilde{P}|_{\mathscr{F}_n}$. 分别表示$P$和$\tilde{P}$到$(\Omega,\mathscr{F}_n)$上的限制. 在第二章,已介绍过测度之间的绝对连续性和相互奇异性. $\tilde{P}\ll P$表示$\tilde{P}$是P-连续的. $\tilde{P}\sim P$表示相互绝对连续或等价. $\tilde{P}\overset{\mathrm{loc}}{\ll}P$表示$\tilde{P}$关于$P$局部绝对连续,即$\tilde{P}_n\ll P_n(\forall\ n\in\mathrm{N})$. $\tilde{P}\overset{\mathrm{loc}}{\sim}P$表示$\tilde{P}$与$P$局部等价,即$\tilde{P}_n\sim P_n(\forall\ n\in\mathrm{N})$. 显然

$$\tilde{P}\ll P\Rightarrow\tilde{P}\overset{\mathrm{loc}}{\ll}P,\tilde{P}\sim P\Rightarrow\tilde{P}\overset{\mathrm{loc}}{\sim}P.$$

但反之不必成立. 这有例可以证明. 现在,一个很自然的问题就是:在什么条件下,逆关系真. 本节将以鞅论为工具来探求这样的条件.

假定$\tilde{P}\overset{\mathrm{loc}}{\ll}P$. 令$z_n=\dfrac{\mathrm{d}\tilde{P}_n}{\mathrm{d}P_n}$,则$z=(z_n,n\in\mathrm{N})$为鞅,且$L^1$-有界. $z_\infty=\overline{\lim}z_n$. 以下符号$E$总表示关于$P$求期望. 按 Doob 鞅收敛定理 4.5.2,有

$$(\mathrm{a.s})-\lim_{n\to\infty}z_n=z_\infty\in L^1.$$

定理 4.9.1(Lebesque 分解定理). 设$\tilde{P}\overset{\mathrm{loc}}{\ll}P$,则

$$\tilde{P}=\nu+\mu,$$

其中$\mu\perp P$,$\nu\ll P$,且$\nu(A)=\int_A z_\infty\mathrm{d}P(\forall\ A\in\mathscr{F})$,

$$\mu(A)=\tilde{P}(A\cap\{z_\infty=\infty\})\ (\forall\ A\in\mathscr{F}).$$

证 引入概率测度$Q=\dfrac{1}{2}(P+\tilde{P})$,$Q_n=\dfrac{1}{2}(P_n+\tilde{P}_n)$. 则$\tilde{P}\ll Q$,$P\ll Q$. 按假设$\tilde{P}_n\ll P_n$,有$P_n\sim Q_n(\forall\ n\in\mathrm{N})$. 按 R.-N. 导数的

性质,$\frac{d\tilde{P}_n}{dQ_n}=\frac{d\tilde{P}_n}{dP_n}\cdot\frac{dP_n}{dQ_n}(\text{mod}Q)$. 令

$$\tilde{u}_\infty=\frac{d\tilde{P}}{dQ},\tilde{u}_\infty=\frac{dP}{dQ},u_n=\frac{d\tilde{P}_n}{dQ_n},u_n=\frac{dP_n}{dQ_n},$$

则

$$Q(\tilde{u}_\infty=0,u_\infty=0)=0(\because \tilde{P}(A)=0=P(A)\Rightarrow Q(A)=0).$$

因此,$\tilde{u}_\infty\cdot u_\infty^{-1}$有定义于$\Omega\backslash\{\tilde{u}_\infty=0,u_\infty=0\}$上. $(\text{mod}Q)$. 显然,$z_n=\frac{\tilde{u}_n}{u_n}(\text{mod}Q)$. 按 Doob 鞅不等式,有

$$Q(\sup_n u_n>c)=\sup_n Q(\max_{0\leqslant k\leqslant n}u_k>c)\leqslant\frac{1}{c}\sup_n E_Q u_n=\frac{1}{c}\xrightarrow[c\to\infty]{}0.$$

$$P(\sup_n u_n>c)\xrightarrow[c\to\infty]{}0(\because P\ll Q)$$

$$\int_{\{u_k>c\}}u_k dQ_k=P(u_k>c)\leqslant P(\max_{0\leqslant l\leqslant k}u_l>c)\leqslant P(\sup_k u_k>c)$$

$$\Rightarrow\lim_{c\to\infty}\sup_k\int_{\{u_k>c\}}u_k dQ\leqslant\lim_{c\to\infty}P(\sup_k u_k>c)=0.$$

这表明:$u.$关于Q为一致可积鞅. 同理可证:$\tilde{u}$关于Q亦为一致可积鞅. 按 Doob-Levy 的鞅表现定理,有

$$\tilde{u}_n=E_Q(\tilde{u}_\infty/\mathscr{F}_n),u_n=E_Q(u_\infty/\mathscr{F}_n)\ (\forall\ n\in\mathrm{N})$$

$$\Rightarrow z_\infty=(\text{a.s})-\lim_{n\to\infty}\frac{\tilde{u}_n}{u_n}=\frac{\tilde{u}_\infty}{u_\infty}\ (\text{mod}Q)$$

$$\Rightarrow z_\infty=\frac{\tilde{u}_\infty}{u_\infty}\ (\text{mod}P\text{ 或 mod}\tilde{P}).$$

现在,令$\nu(A)=\int_A z_\infty dP,\mu(A)=\tilde{P}(A\cap\{z_\infty=\infty\})\ (\forall\ A\in\mathscr{F})$,则

$$\tilde{P}(A)=\nu(A)+\mu(A)\ (\forall\ A\in\mathscr{F}).$$

事实上,设$u_\infty^*=u_\infty^{-1}1_{(u_\infty>0)}$,则

$$\tilde{P}(A)=\int_A\tilde{u}_\infty dQ=\int_A\tilde{u}_\infty u_\infty u_\infty^* dQ+\int_A\tilde{u}_\infty(1-u_\infty\cdot u_\infty^*)dQ$$

$$=\int_A \tilde{u}_\infty u_\infty^* \mathrm{d}P+\int_A (1-u_\infty u_\infty^*)\mathrm{d}\tilde{P}$$

$$=\int_A z_\infty \mathrm{d}P+\tilde{P}(A\cap\{u_\infty=0\}).$$

$$\tilde{P}(A\cap\{u_\infty=0\})=\tilde{P}(A\cap\{u_\infty=0,\tilde{u}_\infty=0\})$$

$$+\tilde{P}(A\cap\{u_\infty=0,\tilde{u}_\infty>0\})$$

$$=\tilde{P}(A\cap\{z_\infty=\infty\})\Rightarrow\tilde{P}=\nu+\mu.$$

最后证明：$\mu\perp P$. 注意：

$$\{z_\infty=\infty\}\subseteq\{u_\infty=0\},P(u_\infty=0)=0.$$

而有 $P(z_\infty=\infty)=0\Rightarrow P(z_\infty<\infty)=1$. 注意，

$$\mu(z_\infty<\infty)=\tilde{P}(\{z_\infty<\infty\}\cap\{z_\infty=\infty\})=0.$$

令 $A=\{z_\infty<\infty\}$，则 $P(A)=1,\mu(A)=0$. 故 $\mu\perp P$.

注1 $z=(z_k,k\in N)$ 在概率 P 下为 L'-有界鞅. $\lim\limits_{n\to\infty}z_n=z_\infty(\mathrm{mod}P)$. 这是在条件 $\tilde{P}\overset{\mathrm{loc}}{\ll}P$ 下得到的. 一般地即使 $P(z_\infty=\infty)=0$，也得不出 $\tilde{P}(z_\infty=\infty)=0$的结果. 若能得出，则 $\tilde{P}\ll P$，可见，z_∞将起特别重要的作用.

注2 在本章 §5.3中给出过一个 L'-有界的非负鞅 η. 它不一致可积. 但有 $\eta_n\xrightarrow{\text{a.s}}\eta_\infty\in L'$. 对$\forall$ $0<\varepsilon<1$，有

$$P(\eta_n\geqslant\varepsilon)=P(\xi_k=1,0\leqslant k\leqslant n)=\left(\frac{1}{2}\right)^{n+1}\xrightarrow[n\to\infty]{}0$$

$$\Rightarrow\eta_n\xrightarrow{P.}0,\Rightarrow\eta_\infty=0\ (\text{a.s}).$$

定义 $\tilde{P}_n$：$\tilde{P}_n(A)=\int_A\eta_n\mathrm{d}P(\forall\ A\in\mathscr{F}_n)$，即 $\eta_n=\dfrac{\mathrm{d}\tilde{P}_n}{\mathrm{d}P_n}$. 按 η 的性质，$\tilde{P}_n$ 为概率测度，且 $\tilde{P}_n\ll P_n(\forall\ n\in\mathrm{N})$. 此外，概率测度族$(\tilde{P}_n,n\in\mathrm{N})$是相容的. 按 K.-连续延拓定理，$\exists$ $(\Omega,\mathscr{F})$ 上的概率测度 $\tilde{P}$，$\ni\tilde{P}|_{\mathscr{F}_n}=\tilde{P}_n$. 按定理4.9.1，有

$$\tilde{P}(A) = \int_A \eta_\infty \mathrm{d}P + \tilde{P}(A \cap \{\eta_\infty = \infty\})$$

$$= \tilde{P}(A \cap \{\eta_\infty = \infty\}) \quad (\forall A \in \mathscr{F}). \Rightarrow \tilde{P} \perp P.$$

然而，$\eta_\infty = 0$ (mod$P$)，$\eta_\infty = \infty$(mod$\tilde{P}$).

定理4.9.2 设 $\tilde{P} \overset{\text{loc}}{\ll} P$，则

$$\tilde{P} \underset{(\perp)}{\ll} P \Leftrightarrow Ez_\infty \underset{(=0)}{=} 1 \Leftrightarrow \tilde{P}(z_\infty < \infty) \underset{(=0)}{=} 1.$$

证 按定理4.9.1，有

$$\tilde{P}(A) = \int_A z_\infty \mathrm{d}P + \tilde{P}(A \cap \{z_\infty = \infty\}) \ (\forall A \in \mathscr{F}). \tag{4.9-1}$$

让 $A = \Omega$，得：$1 = Ez_\infty + \tilde{P}(z_\infty = \infty)$.

若 $\tilde{P} \ll P$，则 $P(z_\infty = \infty) = 0 \Rightarrow \tilde{P}(z_\infty = \infty) = 0. \Rightarrow Ez_\infty = 1$. 反之，若 $Ez_\infty = 1$，则 $\tilde{P}(z_\infty = \infty) = 0$. 按(4.9-1)式，有

$$\tilde{P}(A) = \int_A z_\infty \mathrm{d}P (\forall A \in \mathscr{F}). \Rightarrow \tilde{P} \ll P.$$

仍以(4.9-1)式为依据，便可得命题中相应(·)的结论.

推论4.9.1 若 $\tilde{P} \overset{\text{loc}}{\ll} P$，则 $\tilde{P} \ll P \Leftrightarrow z$ 一致可积(关于 P).

证 若 z 关于 P 一致可积，则 z 为一致可积鞅. 按Doob-Levy定理，

$$z_n = E(z_\infty / \mathscr{F}_n) \Rightarrow Ez_\infty = Ez_n = 1 \Rightarrow \tilde{P} \ll P.$$

现设 $Ez_\infty = 1$，则按定理4.9.2，$\tilde{P} \ll P$. 按(4.9-1)式可得 $\tilde{P}(A) = \int_A z_\infty \mathrm{d}P (\forall A \in \mathscr{F})$. 让 $A \in \mathscr{F}_n$，则

$$\tilde{P}_n(A) = \tilde{P}(A) = \int_A z_\infty \mathrm{d}P = \int_A E(z_\infty | \mathscr{F}_n) \mathrm{d}P_n$$

$$\Rightarrow z_n = \frac{\mathrm{d}\tilde{P}_n}{\mathrm{d}P_n} = E(z_\infty / \mathscr{F}_n) \ (\mathrm{mod}P) \ (\forall n \in \mathrm{N}).$$

$\Rightarrow z$ 一致可积.

注3 按 Doob 鞅不等式(定理4.4.1). 若

$$\sup_n E(z_n \ln^+ z_n) < \infty \text{ 或} \sup_n E z_n^p < \infty (\text{某 } p > 1),$$

则 z 一致可积. 在注4.9.2中出现的鞅 η 经直接验证不是一致可积的. 这里也可采用推论4.9.1,验证 η 不是一致可积的.

定理4.9.3(Kakutani-Dichotomy 定理) 设 $\xi,\tilde{\xi}$ 是定义在 $(\Omega,\mathscr{F},P)$ 上的两上独立的 r.v. 列,且 $\xi,\tilde{\xi}\in A_s(\mathscr{F}.)$,其中 $\mathscr{F}_n = \sigma(\xi_k,\tilde{\xi}_k:0\leqslant k\leqslant n)(\forall\ n\in\mathrm{N})$. $Q,\tilde{Q}$ 分别表示由 ξ 和 $\tilde{\xi}$ 所诱导出的 $(R^\infty,\mathscr{B}^\infty)$ 上的概率测度. $Q_n,\tilde{Q}_n$ 分别为 $Q,\tilde{Q}$ 到 $(R^{n+1},\mathscr{B}^{n+1})$ 上的限制. $P_{\xi_n},P_{\tilde{\xi}_n}$ 分别表示由 $\xi_n,\tilde{\xi}_n$ 可诱导的 $(R,\mathscr{B})$ 上的概率测度.

若 $P_{\tilde{\xi}_n}\ll P_{\xi_n}(\forall\ n\in\mathrm{N})$,则 $\tilde{Q}\ll Q$ 与 $\tilde{Q}\perp Q$ 两者必居其一.

证 由独立性知

$$P_{\tilde{\xi}_n}\ll P_{\xi_n}\Leftrightarrow \tilde{Q}_n\ll Q_n(\forall\ n\in\mathrm{N}), \text{即 } \tilde{Q}\overset{\mathrm{loc}}{\ll} Q.$$

令
$$q_k = \frac{\mathrm{d}P_{\tilde{\xi}_n}}{\mathrm{d}P_{\xi_n}}, \text{则 } z_n = \frac{\mathrm{d}\tilde{Q}_n}{\mathrm{d}Q_n} = \prod_{k=0}^{n} q_k,$$

$$z_\infty = \prod_{n=0}^{\infty} q_n = \lim_{n\to\infty} z_n(\mathrm{mod}Q).$$

令
$$\mathscr{B}_k^\infty = \mathop{\times}_{i=k}^{\infty}\mathscr{B}. \text{ 则 } \bigcap_{k=0}^{\infty}\mathscr{B}_k^\infty \text{ 为 } \mathscr{B}^\infty \text{ 中的尾 } \sigma\text{- 代数}.$$

$$\{z_\infty = 0\} = \{x\in R^\infty: \sum_{k=n}^{\infty}\ln q_k(x_k) = -\infty\}$$

$$= \{x\in R^\infty: \sum_{k=n}^{\infty}\ln q_k(x_k) = -\infty\}\in\mathscr{B}_n^\infty(\forall\ n\in\mathrm{N})$$

$$\Rightarrow \{z_\infty = 0\}\in\bigcap_{k=0}^{\infty}\mathscr{B}_k^\infty.$$

$$\{0 < z_\infty < \infty\} = \{x\in R^\infty: |\sum_{k=0}^{\infty}\ln q_k(x_k)| < \infty\}$$

$$= \{x\in R^\infty: |\sum_{k=n}^{\infty}\ln q_k(x_k)| < \infty\}\in\mathscr{B}_n^\infty(\forall\ n\in\mathrm{N})$$

$$\Rightarrow \{0 < z_\infty < \infty\} \in \bigcap_{k=0}^{\infty} \mathscr{B}_k^\infty.$$

$$\Rightarrow \{z_\infty < \infty\}, \{z_\infty > 0\} \in \bigcap_{k=0}^{\infty} \mathscr{B}_k^\infty.$$

按 $K.$-0-1律，$\tilde{Q}(z_\infty < \infty) = 0$或1. 按定理4.9.2，即得结论.

引理4.9.1 设

$$u(x) = x \cdot 1_{\{|x| \leqslant 1\}}(x) + \text{sign}(x) 1_{\{|x|>1\}}(x) \ (\forall\, x \in \mathrm{R}),$$

则∃ 常数$0 < A < B < \infty$，使得

$$A(1 - \sqrt{x})^2 \leqslant x u(\ln x) + x u^2(\ln x) + 1 - x$$
$$\leqslant B \cdot (1 - \sqrt{x})^2 (\forall\, x \geqslant 0).$$

证 令 $f(x) = x \cdot u(\ln x) + x u^2(\ln x) + 1 - x (x > 0)$. 首先考虑 $x \geqslant \mathrm{e}$ 的情形. 让 $f_1(x) = \dfrac{f(x)}{(1-\sqrt{x})^2} = \dfrac{1+x}{(1-\sqrt{x})^2}$，则 $f_1(x) \downarrow$ 当 $x \uparrow$ 时($x \geqslant \mathrm{e}$). 因此，$1 = f_1(\infty) \leqslant f_1(x) \leqslant f_1(\mathrm{e})$. 于是，当 $x \geqslant \mathrm{e}$ 时，选取 A、B 满足要求：$A \leqslant 1, B \geqslant f_1(\mathrm{e})$. 其次考虑$0 \leqslant x \leqslant \mathrm{e}^{-1}$的情形. 这时 $f_1(x) = \dfrac{f(x)}{(1-\sqrt{x})^2} = \dfrac{1+\sqrt{x}}{1-\sqrt{x}}$. 显然，$f_1(x) \uparrow$ 当$0 \leqslant x \leqslant \mathrm{e}^{-1}$且 $x \uparrow$ 时. 于是，$1 = f_1(0) \leqslant f_1(x) \leqslant f_1(\mathrm{e}^{-1})$. 这时，应选取 A, B 满足要求 $A \leqslant 1, B \geqslant f_1(\mathrm{e}^{-1})$. 再次考虑$1 \leqslant x \leqslant \mathrm{e}$ 的情形. 让 $f_2(x) = f(x) - c(1 - \sqrt{x})^2$，则

$$f'_2(x) = 3\ln x + (\ln x)^2 - c + \frac{c}{\sqrt{x}},$$

$$f''_2(x) = \frac{3}{x} + \frac{2\ln x}{x} - \frac{c}{2} \frac{1}{x\sqrt{x}}.$$

不难看出：当 $c \leqslant 6$时，$f_2(x) \geqslant 0$于 $x \in [1, \mathrm{e}]$上；当 $c > 2\sqrt{\mathrm{e}}(2\mathrm{e}+1)$时，$f_2(x) \leqslant 0$于 $x \in [1, \mathrm{e}]$上. 于是，当 $x \in [1, \mathrm{e}]$时，选取的 A, B 应满足要求 $A \leqslant 6, B \geqslant 28 > 2\sqrt{\mathrm{e}}(2\mathrm{e}+1)$. 最后考虑 $\mathrm{e}^{-1} \leqslant x \leqslant 1$的情形. 仍考虑函数 $f_2(x)$. 不难看出：当 $c \leqslant 2$时，$f_2(x) \geqslant 0$于 $x \in [\mathrm{e}^{-1}, 1]$上；当 $c \geqslant 10 > 2(5 - 2\mathrm{e}^{-1})$时，$f_2(x) \leqslant 0$于 $x \in [\mathrm{e}^{-1}, 1]$上.

于是，当 $x\in[e^{-1},1]$ 时，选取的 A、B 应满足要求：$A\leqslant 2$，$B\geqslant 10$.

综上所述，可选取 A、B 为 $A=1$，$B=28$.

定理4.9.4 设 $\tilde{P}\overset{loc}{\ll}P$，$\alpha_n=z_nz_{n-1}^*$，$z_{n-1}^*=\dfrac{1}{z_{n-1}}1_{\{z_{n-1}>0\}}(n\geqslant 1)$. $\mathscr{F}_0=\{\varnothing,\Omega\}$，$z_0=1$. 则

$$\tilde{P}\underset{(\perp)}{\ll}P\Leftrightarrow\tilde{P}\left(\sum_n(1-E(\sqrt{\alpha_n}/\mathscr{F}_{n-1}))<\infty\right)\underset{(=0)}{=}1.$$

证 按定理4.9.2，$\tilde{P}\ll P\Leftrightarrow\tilde{P}(z_\infty<\infty)=1$. 下面讨论 $\{z_\infty<\infty\}$ 的结构. 按假设，

$$\tilde{P}_n(z_n=0)=0=\tilde{P}_n(z_n=\infty)\ (\forall\ n\in\mathrm{N})$$

$$\Rightarrow z_n=\prod_{k=1}^n z_k\cdot z_{k-1}^*=\prod_{k=1}^n\alpha_k=\exp\left(\sum_{k=1}^n\ln\alpha_k\right)(\mathrm{mod}\tilde{P}\text{ 或 }\mathrm{mod}P).$$

按定理4.9.1，有

$$\tilde{P}(z_\infty=0)=0$$

$$\Rightarrow\{z_\infty<\infty\}=\{0<\lim_{n\to\infty}z_n<\infty\}$$

$$=\{-\infty<\lim_{n\to\infty}\sum_{k=1}^n\ln\alpha_k<+\infty\}$$

$$=\{\lim_{n\to\infty}|\sum_{k=1}^n u(\ln\alpha_k)|<\infty\}(\mathrm{mod}\tilde{P}\text{ 或 }\mathrm{mod}P),$$

其中 $u(x)$ 按引理4.9.1中定义(注意：$|\ln\alpha_k|\to 0$ 当 $k\uparrow\infty$ 时).

现在，设 η 为任意有界的 r.v.，则对 $\forall\ A\in\mathscr{F}_{n-1}$，有

$$\int_A\tilde{E}(\eta z_{n-1}/\mathscr{F}_{n-1})\mathrm{d}\tilde{P}=\int_A\eta z_{n-1}\mathrm{d}\tilde{P}=\int_A\eta z_{n-1}\cdot z_n\mathrm{d}P$$

$$=\int_A z_{n-1}E(\eta z_n/\mathscr{F}_{n-1})\mathrm{d}P=\int_A E(\eta z_n/\mathscr{F}_{n-1})\mathrm{d}\tilde{P}.$$

$$\int_A\eta z_n\mathrm{d}P=\int_A\eta\mathrm{d}\tilde{P}=\int_A\tilde{E}(\eta/\mathscr{F}_{n-1})\mathrm{d}\tilde{P}=\int_A z_{n-1}\tilde{E}(\eta/\mathscr{F}_{n-1})\mathrm{d}P$$

$$\Rightarrow z_{n-1}\tilde{E}(\eta/\mathscr{F}_{n-1})=E(\eta z_n/\mathscr{F}_{n-1})(\mathrm{mod}P\text{ 或 }\mathrm{mod}\tilde{P}).$$

(这里 $\tilde{E}$ 表示关于 $\tilde{P}$ 求期望).

令 $\eta=z_{n-1}^*$，则$1=E(\alpha_n/\mathscr{F}_{n-1})(\mathrm{mod}\tilde{P})$

令 $\eta=u(\ln\alpha_n)$，则

$$\tilde{E}(u(\ln\alpha_n)/\mathscr{F}_{n-1})=E(\alpha_n\cdot u(\ln\alpha_n)/\mathscr{F}_{n-1})(\mathrm{mod}\tilde{P}).$$

注意，当 x 在1附近时，$xu(\ln x)\geqslant x-1$. 于是，当 n 足够大时，有

$$\tilde{E}(u(\ln\alpha_n)/\mathscr{F}_{n-1})=E(\alpha_n u(\ln\alpha_n)/\mathscr{F}_{n-1})$$
$$\geqslant E(\alpha_n-1/\mathscr{F}_{n-1})=0\ (\mathrm{mod}\tilde{P}).$$

令 $\xi_n=\sum_{k=1}^{n}u(\ln\alpha_k)$，则 ξ 关于 $\tilde{P}$ 为下鞅，且当 n 足够大时，$|\Delta\xi_n|\leqslant 2$. 按定理4.8.6，得

$$\{\lim_{n\to\infty}|\sum_{k=1}^{n}u(\ln\alpha_k)|<\infty\}=\{\xi_n\to\}$$
$$=\{\sum_{k=1}^{\infty}\tilde{E}(u(\ln\alpha_k)+u^2(\ln\alpha_k)/\mathscr{F}_{k-1})<\infty\}$$
$$=\{\sum_{k=1}^{\infty}E(\alpha_k(u(\ln\alpha_k)+u^2(\ln\alpha_k))/\mathscr{F}_{-1})<\infty\}(\mathrm{mod}\tilde{P}).$$

(最后一步是取 $\eta=u(\ln\alpha_k)+u^2(\ln\alpha_k)$得到的). 注意

$$E((1-\sqrt{\alpha_n})^2/\mathscr{F}_{n-1})=2E(1-\sqrt{\alpha_n}/\mathscr{F}_{n-1})(\mathrm{mod}\tilde{P}).$$

并利用引理4.9.1，便可推出：

$$\sum_n E(\alpha_n(u(\ln\alpha_n)+u^2(\ln\alpha_n))/\mathscr{F}_{n-1})<\infty$$
$$\Leftrightarrow\sum_n E(1-\sqrt{\alpha_n}/\mathscr{F}_{n-1})<\infty(\mathrm{mod}\tilde{P}).$$
$$\Rightarrow\{z_\infty<\infty\}=\{\sum_n(1-E(\sqrt{\alpha_n}/\mathscr{F}_{n-1}))<\infty\}(\mathrm{mod}\tilde{P}).$$

推论4.9.2 设 $\tilde{P}\overset{\mathrm{loc}}{\ll}P$，$\alpha_n$ 与 $\mathscr{F}_{n-1}$关于 P 和 $\tilde{P}$ 独立，则 $\tilde{P}\underset{(\perp)}{\ll}P\Leftrightarrow\sum_n(1-E\sqrt{\alpha_n})\underset{(=)}{<}\infty.$

注4 按此推论，Kakutani 定理可表述为

$$\tilde{Q}\underset{(\perp)}{\ll}Q\Leftrightarrow\sum_n(1-E\sqrt{q_n})\underset{(=)}{<}\infty.$$

推论4.9.3 若 $\tilde{P}\overset{\text{loc}}{\ll}P$，则

$$\tilde{P}(\sum_n E(\alpha_n\cdot\ln\alpha_n/\mathscr{F}_{n-1})<\infty)=1\Rightarrow\tilde{P}\ll P.$$

证 在定理4.9.4之证中得到过

$$E(1-\alpha_n/\mathscr{F}_{n-1})=0\ (\bmod\tilde{P}).$$

由此可知，对$\forall\ C\in\mathrm{R}$，有

$$\tilde{P}(\sum_n E(\alpha_n\cdot\ln\alpha_n/\mathscr{F}_{n-1})<\infty)$$

$$=\tilde{P}(\sum_n E(\alpha_n\cdot\ln\alpha_n+c(1-\alpha_n)/\mathscr{F}_{n-1})<\infty);$$

注意
$$\tilde{P}(\sum_n E(\alpha_n\cdot\ln\alpha_n+1-\alpha_n/\mathscr{F}_{n-1})<\infty)$$

$$\leqslant\tilde{P}(\sum_n E(1-\sqrt{\alpha_n}/\mathscr{F}_{n-1})<\infty).$$

于是，按定理4.9.4，并取$c=1$，便得本推论的结果.

推论4.9.4 若 $\tilde{P}\overset{\text{loc}}{\ll}P$，则

$$\tilde{P}\underset{(\perp)}{\ll}P\Leftrightarrow\tilde{P}(\sum_n|\ln E(\sqrt{\alpha_n}/\mathscr{F}_{n-1})|<\infty)=\underset{(0)}{1}.$$

证 对$\forall\ x\in(0,1]$，有$-x\ln x\leqslant 1-x\leqslant-\ln x$. 因此，若$0<x_n\leqslant 1$，且$\sum_n(1-x_n)<\infty$，则$-\sum_n x_n\ln x_n<\infty$，且 $x_n\uparrow 1$. 从而，$\sum_n|\ln x_n|<\infty$. 反之，由此不等式可推出：当$0<x_n<1$时，$\sum_n(1-x_n)<\infty$. 从而，$|\sum_n x_n\ln x_n|<\infty$. 故当$0<x_n\leqslant 1$时，有

$$\sum_n(1-x_n)<\infty\Leftrightarrow|\sum_n x_n\ln x_n|<\infty\Leftrightarrow\sum_n|\ln x_n|<\infty.$$

令 $x_n=E(\sqrt{\alpha_n}/\mathscr{F}_{n-1})$，则

$$x_n^2\leqslant E(\alpha_n/\mathscr{F}_{n-1})=1(\bmod\tilde{P})，\text{即 }0<x_n\leqslant 1(\bmod\tilde{P})$$

$$\Rightarrow\tilde{P}(\sum_n(1-E(\sqrt{\alpha_n}/\mathscr{F}_{n-1}))<\infty)$$

$$= 1 \Leftrightarrow \tilde{P}\left(\sum_n |\ln E(\sqrt{\alpha_n}/\mathscr{F}_{n-1})| < \infty\right) = 1.$$

故此推论便转化成定理4.9.4.

推论4.9.5 设 $\tilde{P} \overset{\text{loc}}{\ll} P$，且∃ $A, B: 0 \leqslant A < 1, 0 < B$，使得

$$P(1 - A \leqslant \alpha_n \leqslant 1 + B) = 1 \ (\forall\, n \geqslant 1),$$

$$\tilde{P} \underset{(\perp)}{\ll} P \Leftrightarrow \tilde{P}\left(\sum_n E((1 - \alpha_n)^2/\mathscr{F}_{n-1}) < 1\right) = \underset{(0)}{1}.$$

证 事实上，若$1-A \leqslant x \leqslant 1+B$，则∃ $c_i > 0 (i=1,2)$，使得

$$c_1(1 - x)^2 \leqslant (1 - \sqrt{x})^2 \leqslant c_2(1 - x)^2. \qquad (4.9\text{-}2)$$

令 $f(x) = \left(\dfrac{1-\sqrt{x}}{1-x}\right)^2 = \dfrac{1}{(1+\sqrt{x})^2}$. 则

$$(1 + \sqrt{1 + B})^{-2} \leqslant f(x) \leqslant (1 + \sqrt{1 - A})^{-2}$$

$$(\forall\, x \in [1 - A, 1 + B]).$$

取 $c_1 = (1+\sqrt{1+B})^{-2}$，$c_2 = (1+\sqrt{1-A})^{-2}$，则(4.9-2)式对任意 $x \in [1-A, 1+B]$成立. 注意

$$E((1 - \sqrt{\alpha_n})^2/\mathscr{F}_{n-1}) = 2E(1 - \sqrt{\alpha_n}/\mathscr{F}_{n-1}) (\text{mod}\tilde{P}).$$

故此推论被转化成定理4.9.4.

下面利用前面所得到的结果讨论 Hajek-Feldman-Dichotomy 问题：设 $P, \tilde{P}$ 为$(R^\infty, \mathscr{B}^\infty)$上的概率测度. $P_n, \tilde{P}_n$ 分别表示 $P, \tilde{P}$ 到$(R^n, \mathscr{B}^n)$上的限制. E、$\tilde{E}$ 分别表示关于 $P, \tilde{P}$ 求期望. H.-F.-D. 问题将研究Gauss 测度族之间的等价关系.

命题4.9.1 设 $x = (x_n, n \geqslant 1)$为 R^∞上的随机点. $P = (P_n, n \geqslant 1)$为 Gauss 概率族，则

$$E(x_n/\mathscr{B}^{n-1}) = E(x_n/x_1, \cdots, x_{n-1}) = \sum_{k=1}^{n-1} c_k x_k = a_{n-1}(x) (n \geqslant 2),$$

其中 c_k，$1 \leqslant k \leqslant n-1$，为常数，且

存在独立同分布的 Gauss r.v. 列 $\varepsilon = (\varepsilon_n, n \geqslant 1)$ $(\varepsilon_n \sim N(0, 1))$，

∋ $x_n = a_{n-1}(x) + b_{n-1} \cdot \varepsilon_n$ (b_{n-1}为常数).

证 这可由 Gram-Schmidt 方法推出(参看[1]. ch. 2. § 11).

设 $\tilde{P}=(\tilde{P}_n,n\geqslant1)$ 为一个 Gauss 测度族,则相应地可得 $\tilde{b}_{n-1}$ 及 $\tilde{a}_{n-1}(x)=\tilde{E}(x_n/\mathscr{B}^{n-1})=\sum_{k=1}^{n-1}\tilde{c}_k x_k$. 如果 $\tilde{P}\overset{\text{loc}}{\ll}P$,则当 $b_{n-1}=0$ 或 $\tilde{b}_{n-1}=0$ 时,必有 $\tilde{a}_{n-1}(x)=a_{n-1}(x)$. $(\text{mod}P)$. 由此知:$\tilde{a}_k(x)=a_k(x)$ $(\text{mod}P)(1\leqslant k\leqslant n-1)$. 可以证明:

$$\tilde{a}_k(x)=\frac{\mathrm{d}\tilde{P}_k}{\mathrm{d}P_k}(x)\cdot a_k(x)(\text{mod}P)(1\leqslant k\leqslant n-1)$$

$$\Rightarrow z_k=\frac{\mathrm{d}\tilde{P}_k}{\mathrm{d}P_k}=1(\text{mod}P)(1\leqslant k\leqslant n-1).$$

由 $\tilde{P}\overset{\text{loc}}{\sim}P$ 可知:$z_k=1(\text{mod}\tilde{P})$. 这表明:$\tilde{P}_k=P_k(1\leqslant k\leqslant n-1)$. 因此,以下总假定:$b_n^2>0,\tilde{b}_n^2>0$. $(\forall\ n\geqslant1)$.

命题4. 9. 2 设 $P,\tilde{P}$ 为 $(R^\infty,\mathscr{B}^\infty)$ 上的 Gauss 概率测度,且 $\tilde{P}\overset{\text{loc}}{\sim}P$. 则

$$\alpha_n=d_{n-1}^{-1}\cdot\exp\left\{-\frac{1}{2}\left(\frac{(x_n-\tilde{a}_{n-1}(x))^2}{\tilde{b}_{n-1}^2}-\frac{(x_n-a_{n-1}(x))^2}{b_{n-1}^2}\right)\right\},$$

其中 $a_0(x)=Ex_1,\tilde{a}_0(x)=\tilde{E}x_1,b_0^2=Vx_1,\tilde{b}_0^2=\tilde{V}x_1$,

$$d_n=\left|\frac{\tilde{b}_n}{b_n}\right|,\alpha_n=z_n\cdot z_{n-1}^{*}(z\text{ 之定义见定理 4. 9. 4}).$$

证 设 λ 为 $(\mathrm{R},\mathscr{B})$ 上的 Lebesque 测度. $P_{n|n-1}$ 表示 x_n 在条件 $\mathscr{B}^{n-1}$ 下的条件概率,即

$$P_{n|n-1}(B)=P(x_n\in B/\mathscr{B}^{n-1})(\forall\ B\in\mathscr{B}).$$

显然,由 Gauss 性,有 $\frac{\mathrm{d}P_n}{\mathrm{d}P_{n-1}}=\frac{\mathrm{d}P_{n|n-1}}{\mathrm{d}\lambda}$. 注意

$$P(x_n\in B/\mathscr{B}^{n-1})=P(a_{n-1}(x)+b_{n-1}\varepsilon_n\in B/\mathscr{B}^{n-1})(\forall\ B\in\mathscr{B})$$

$$\Rightarrow\frac{\mathrm{d}P_{n|n-1}}{\mathrm{d}\lambda}=\frac{1}{\sqrt{2\pi}\,|b_{n-1}|}\exp\left\{-\frac{1}{2b_{n-1}^2}(x_n-a_{n-1}(x))^2\right\}$$

$$\Rightarrow\frac{\mathrm{d}P_n}{\mathrm{d}P_{n-1}}=\frac{1}{\sqrt{2\pi}\,|b_{n-1}|}\exp\left\{-\frac{1}{2b_{n-1}^2}(x_n-a_{n-1}(x))^2\right\}.$$

同理可得

$$\frac{\mathrm{d}\widetilde{P}_n}{\mathrm{d}\widetilde{P}_{n-1}}=\frac{1}{\sqrt{2\pi}\,|\widetilde{b}_{n-1}|}\exp\left\{-\frac{1}{2\widetilde{b}_{n-1}^2}(x_n-\widetilde{a}_{n-1}(x))^2\right\}$$

$$\Rightarrow\alpha_n(x)=z_nz_{n-1}^*=\frac{\mathrm{d}\widetilde{P}_n}{\mathrm{d}P_n}\cdot\frac{\mathrm{d}P_{n-1}}{\mathrm{d}\widetilde{P}_{n-1}}=\frac{\mathrm{d}\widetilde{P}_n}{\mathrm{d}\widetilde{P}_{n-1}}\cdot\frac{\mathrm{d}P_{n-1}}{\mathrm{d}P_n}$$

$$=d_{n-1}^{-1}\cdot\exp\left\{-\frac{1}{2}\widetilde{b}_{n-1}^{-2}(x_n-\widetilde{a}_{n-1}(x))^2-b_{n-1}^{-2}(x_n-a_{n-1}(x))^2)\right\}.$$

命题4.9.3 设 $\widetilde{P}\overset{\mathrm{loc}}{\sim}P.P$，$P$ 为 $(R^\infty,\mathscr{B}^\infty)$ 上的 Gauss 概率测度，则

$$E(\sqrt{\alpha_n}/\mathscr{B}^{n-1})=\left(\frac{2d_{n-1}}{1+d_{n-1}^2}\right)^{\frac{1}{2}}$$

$$\cdot\exp\left\{-\frac{1}{4}\frac{d_{n-1}^2}{1+d_{n-1}^2}\frac{(a_{n-1}(x)-\widetilde{a}_{n-1}(x))^2}{\widetilde{b}_{n-1}^2}\right\}.$$

即
$$\ln E(\sqrt{\alpha_n}/\mathscr{B}^{n-1})=\frac{1}{2}\ln\left(\frac{2d_{n-1}}{1+d_{n-1}^2}\right)$$

$$-\frac{1}{4}\frac{d_{n-1}^2}{1+d_{n-1}^2}\frac{(a_{n-1}(x)-\widetilde{a}_{n-1}(x))^2}{\widetilde{b}_{n-1}^2}.$$

证 令 $\varphi_{n-1}(x)=a_{n-1}(x)-\widetilde{a}_{n-1}(x)$，$\psi_{n-1}(x)=\frac{1}{b_{n-1}}\varphi_{n-1}(x)$，有

$$f(x,t)=\exp\left\{\frac{1}{4}t^2-\frac{1}{4\widetilde{b}_{n-1}^2}(b_{n-1}t+\varphi_{n-1}(x))^2\right\}\cdot\exp\left\{-\frac{t^2}{2}\right\}$$

$$=\exp\left\{-\frac{t^2}{4}-\frac{1}{4d_{n-1}^2}(t+\psi_{n-1}(x))^2\right\}.$$

则

$$E(\sqrt{\alpha_n}/\mathscr{B}^{n-1})=E_{n|n-1}$$

$$\cdot\left\{d_{n-1}^{-\frac{1}{2}}\exp\left\{\frac{(x_n-a_{n-1}(x))^2}{4b_{n-1}^2}-\frac{(x_n-\widetilde{a}_{n-1}(x))^2}{4\widetilde{b}_{n-1}^2}\right\}\right\}$$

$$= E\left\{ d_n^{-\frac{1}{4}} \exp\left\{ \frac{1}{4}\varepsilon_n^2 - \frac{1}{4\widetilde{b}_{n-1}^2}(b_{n-1}\varepsilon_n + \varphi_{n-1}(x))^2 \right\} \Big/ \mathscr{B}^{n-1} \right\}$$

$$= d_n^{-\frac{1}{4}} \int_R \frac{1}{\sqrt{2\pi}} f(x,t) \mathrm{d}t.$$

计算上式右边的积分便可得命题中所要求的表示.

命题4.9.4 $d=(d_n, n\geqslant 1)$具有下列特性.

(i) $0 \leqslant \dfrac{2d_{n-1}}{1+d_{n-1}^2} \leqslant 1$;

(ii) $\sum\limits_n \ln\left(\dfrac{1+d_{n-1}^2}{2d_{n-1}}\right)$与$\sum\limits_n (d_{n-1}^2-1)^2$具有相同的收敛性.

证 令 $g(t)=\dfrac{2t}{1+t^2}$,则 $g'(t)=\dfrac{2(1-t^2)}{(1+t^2)^2}$. 因此,$t=1$为 $g(t)$ 的极大值点. 于是,$g(0)\leqslant g(t)\leqslant g(1)=1(\forall\ t\geqslant 0)$. $\Rightarrow$(i).

现在,令 $f_c(t)=\ln\dfrac{1+t^2}{2t}-c(t^2-1)^2$,则

$$f'_c(t) = \frac{t^2-1}{t(1+t^2)} - 4ct(t^2-1).$$

按(i),仅需考虑 $t\geqslant 0$的情形. (ii)中两个级数收敛的必要条件是 $d_n \to 1$. 因此,仅需在 $t=1$附近改变函数 $f_c(t)$. 若选 c,使得$\dfrac{1}{2}-4c>0$,则 $f_c(t)$以 $t=1$为极小值点. 若选 c,使得$\dfrac{1}{2}-4c<0$,则 $f_c(t)$以 $t=1$为极大值点. 于是,选$0<c_1<\dfrac{1}{8}$,则$\exists\ B>0$,$\ni f_{c_1}(t)\geqslant 0(|t-1|\leqslant B)$. 选 $c_2>\dfrac{1}{8}$,则$\exists\ A>0$,$\ni f_{c_2}(t)\leqslant 0(|t-1|\leqslant A)$. 故

$$c_1(t^2-1)^2 \leqslant \ln\frac{1+t^2}{2t} \leqslant c_2(t^2-1)^2 (1-A\vee B \leqslant t \leqslant 1+A\wedge B).$$

利用此关系便可得结论(ii).

定理4.9.5(Hajek-Feldman-Dichotomy 定理) 设 $\xi=(\xi_n, n\geqslant 1)$, $\widetilde{\xi}=(\widetilde{\xi}_n, n\geqslant 1)$为两个 Gauss 序列. P, $\widetilde{P}$分别表示由 ξ, $\widetilde{\xi}$ 诱导的$(R^\infty, \mathscr{B}^\infty)$上的Gauss 测度.

若 $\widetilde{P}\overset{\text{loc}}{\sim}P$，则

$$\widetilde{P}\underset{(\perp)}{\sim}P\Leftrightarrow\sum_n\left[\left(\left|\frac{\widetilde{b}_n}{b_n}\right|^2-1\right)^2+\widetilde{E}\left(\frac{\varphi_n^2(x)}{\widetilde{b}_n^2}\right)\right]\underset{(=)}{\leqslant}\infty,$$

其中 $\varphi_n(x)=a_n(x)-\widetilde{a}_n(x)$.

证　按命题4.9.3，4.9.4及推论4.9.4，有

$$\widetilde{P}\underset{(\perp)}{\ll}P\Leftrightarrow\widetilde{P}\left\{\sum_n(d_n^2-1)^2+\frac{\varphi_n^2(x)}{\widetilde{b}_n^2}<\infty\right\}=\underset{(0)}{1}.$$

因此，若能证明

$$\widetilde{P}\left\{\sum_n\frac{\varphi_n^2(x)}{\widetilde{b}_n^2}<\infty\right\}=1\Leftrightarrow\sum_n\widetilde{E}\left\{\frac{\varphi_n^2(x)}{\widetilde{b}_n^2}\right\}<\infty,$$

则由 P、$\widetilde{P}$ 处于对称位置必得定理真. 此等价关系的成立是如下命题的结论.

命题4.9.5　设 $\beta=(\beta_n,n\geqslant1)$ 为定义在 $(\Omega,\mathscr{F},P)$ 上的 Gauss 序列，则

$$P(\sum_n\beta_n^2<\infty)=1\Leftrightarrow\sum_n E\beta_n^2<\infty.$$

证　若 $\sum\limits_n E\beta_n^2<\infty$，则按单调收敛定理，有

$$E(\sum_n\beta_n^2)=\sum_n E\beta_n^2<\infty\Rightarrow P(\sum_n\beta_n^2<\infty)=1.$$

反之，设 $P(\sum\limits_n\beta_n^2<\infty)=1$. 首先假定 β_n 具有0均值($\forall\ n\geqslant 1$). 令 $\eta=\sum\limits_{k=1}^{\infty}\beta_k^2$，则$\exists\ a>0$，$\ni P_\eta((-\infty,a])=P(\eta\leqslant a)>0$

$$\Rightarrow E\exp(-\eta)=\int_0^\infty e^{-t}P_\eta(dt)$$

$$\geqslant\int_0^a e^{-t}P_\eta(dt)\geqslant e^{-a}\cdot P_\eta([0,a])=e^{-a}P(\eta\leqslant a)$$

$$\Rightarrow(E\exp(-\eta))^{-2}\leqslant(e^{-a}P(\eta\leqslant a))^{-2}<\infty;$$

即
$$(E\exp(-\sum_n\beta_n^2))^{-2}<\infty.$$

现在，对$\forall\ n\geqslant 1$，构作独立的 Gauss r. v. 族：$\beta_{k,n}$，$1\leqslant k\leqslant n$，使得

$$\sum_{k=1}^{n}\beta_k^2=\sum_{k=1}^{r}\beta_{k,n}^2(1\leqslant r\leqslant n)，且 E\beta_{k,n}=0.$$

令 $\lambda_{k,n}=E\beta_{k,n}^2$，则 $\beta_{k,n}\sim N(0,\lambda_{k,n})$. 于是

$$\begin{aligned}E\exp\left\{-\sum_{k=1}^{r}\beta_{k,n}^2\right\}&=\prod_{k=1}^{r}E\exp(-\beta_{k,n}^2)\\&=\prod_{k=1}^{r}\int_R\frac{1}{\sqrt{2\pi}}\exp\left\{-t^2-\frac{t^2}{2\lambda_{k,n}}\right\}\frac{\mathrm{d}t}{\sqrt{\lambda_{k,n}}}\\&=\prod_{k=1}^{r}(1+2\lambda_{k,n})^{-\frac{1}{2}}\end{aligned}$$

$$\Rightarrow\left(E\exp\left\{-\sum_{k=1}^{r}\beta_{k,n}^2\right\}\right)^{-2}=\prod_{k=1}^{r}(1+2\lambda_{k,n})=1+2\sum_{k=1}^{r}\lambda_{k,n}+\cdots$$

$$\begin{aligned}\Rightarrow E\left(\sum_{k=1}^{n}\beta_k^2\right)&=E\left(\sum_{k=1}^{r}\beta_{k,n}^2\right)=\sum_{k=1}^{r}\lambda_{k,n}\leqslant\prod_{k=1}^{r}(1+2\lambda_{k,n})\\&\leqslant\left(E\exp\left\{-\sum_{k=1}^{n}\beta_k^2\right\}\right)^{-2}\\&\Rightarrow E\left(\sum_{n}\beta_n^2\right)\leqslant\left(E\exp\left\{-\sum_{n}\beta_n^2\right\}\right)^{-2}\\&\Rightarrow E\left(\sum_{n}\beta_n^2\right)<\infty.\end{aligned}$$

现在，考虑 $E\beta_n\neq 0.$ $(\forall\ n\geqslant 1)$的情形. 设 $\widetilde{\beta}=(\widetilde{\beta}_n,n\geqslant 1)$为一独立于 β 的 Gauss 序列，且具有与 β 相同的分布. 令 $\mu_n=\beta_n-\widetilde{\beta}_n$ $(\forall\ n\geqslant 1)$，则 $E\mu_n=0$，且 μ 为 Gauss 序列. 按前面的证明，有

$$P\left(\sum_{n}\beta_n^2<\infty\right)=1\Leftrightarrow P\left(\sum_{n}(\beta_n-\widetilde{\beta}_n)^2<\infty\right)=1\Leftrightarrow E\left(\sum_{n}\beta_n^2\right)<\infty.$$

剩下要证明的是：

$$E\left(\sum_{n}\mu_n^2\right)<\infty\Leftrightarrow E\left(\sum_{n}\beta_n^2\right)<\infty.$$

按假定，有

$$2\sum_n E(\beta_n - E\beta_n)^2 = \sum_n E\mu_n^2.$$

$$(E\beta_n)^2 \leqslant (E\beta_n)^2 + (2\beta_n - E\beta_n)^2 = 2\beta_n^2 + 2(\beta_n - E\beta_n)^2;$$

$$\sum_n (E\beta_n)^2 \leqslant 2\sum_n \beta_n^2 + 2\sum_n (\beta_n - E\beta_n)^2.$$

因此，若 $P(\sum_n \beta_n^2<\infty)=1$，则 $\sum_n E\mu_n^2<\infty$. 从而

$$\sum_n E(\beta_n - E\beta_n)^2 < \infty \Rightarrow \sum_n (E\beta_n)^2 < \infty$$

$$\Rightarrow E(\sum_n \beta_n^2) = \sum_n E\beta_n^2 = \sum_n (E\beta_n)^2 + \sum_n E(\beta_n - E\beta_n)^2 < \infty.$$

反之，若 $E(\sum_n \beta_n^2)<\infty$，则

$$E(\sum_n \mu_n^2) = \frac{1}{2}\sum_n E(\beta_n - E\beta_n)^2 \leqslant E(\sum_n \beta_n^2) < \infty.$$

上节讨论过一般性的参数估计. 现在，利用 H-F-D 定理来研究噪声为 Gauss 情形的参数估计.

定理4.9.6 设 $\xi=(\xi_n,n\geqslant1)$为独立的 Gauss 序列，且 $E\xi_n=0$，$v_n=E\xi_n^2>0(n\geqslant1)$，$\xi_0=0$，则估计 $\hat{\theta}=(\hat{\theta}_n,n\geqslant1)$强相容的充要条件是$\sum_{n=0}^{\infty}\frac{v_n}{v_{n+1}}=\infty$.

证 设$\sum_n \frac{v_n}{v_{n+1}}=\infty$. 考虑的参数系统为

$$\eta_{n+1} = \theta\eta_n + \xi_{n+1}(\forall\, n \in \mathrm{N}).$$

其中 η_0为独立于 ξ 的 Gauss 变量. P_θ 为 η 所诱导的$(R^\infty,\mathscr{B}^\infty)$上的概率分布. E_θ 表示关于 P_θ 的期望算子. 按上一节的讨论，有如下结果：

$$\hat{\theta}_n = \theta + \frac{\gamma_n}{\alpha_n},$$

其中 $\gamma_n=\sum_{k=0}^{n-1}\frac{\eta_k\cdot\xi_{k+1}}{v_{k+1}}$，$\alpha_n=\sum_{k=0}^{n-1}\frac{\eta_k^2}{v_{k+1}}=\langle\gamma\rangle_n(n\geqslant1)$，$\gamma_0=0=\alpha_0$.

按命题4.9.5及 η 的 Gauss 性，可得

$$P_\theta(\langle\gamma\rangle_\infty=\infty)=1\Leftrightarrow E_\theta\langle\gamma\rangle_\infty=\infty\Leftrightarrow\sum_k\frac{E_\theta\eta_k^2}{v_{k+1}}=\infty.$$

$$E_\theta\eta_k^2=E_\theta(\theta\eta_{k-1}+\xi_k)^2=E_\theta(\sum_{l=0}^{k-1}\theta^l\xi_{k-l}+\theta^k\eta_0)^2$$

$$=\sum_{l=0}^{k-1}\theta^{2l}v_{k-l}+\theta^{2k}E\eta_0^2.$$

$$\sum_{k=1}^\infty\frac{E_\theta\eta_k^2}{v_{k+1}}=\sum_{k=1}^\infty\frac{1}{v_{k+1}}(\sum_{l=0}^{k-1}\theta^{2l}\cdot v_{k-l}+\theta^{2k}E\eta_0^2)$$

$$=\sum_{k=1}^\infty\frac{1}{v_{k+1}}(v_k+\sum_{l=1}^{k}\theta^{2l}v_{k-l})(\text{令 } v_0=E\eta_0^2).$$

$$=\sum_{k=1}^\infty\frac{v_k}{v_{k+1}}+\sum_{l=1}^\infty\sum_{k=l}^\infty\frac{v_{k-l}}{v_{k+1}}\theta^{2l}=\infty.$$

按推论4.9.4，对每个θ，估计$\hat{\theta}$都是强相容的.

现在，假定：$\hat{\theta}$是强相容的. P_θ如前面定义，则$P_\theta(\hat{\theta}_n\to\theta)=1$，若$\theta_1\neq\theta_2$，则必有$P_{\theta_1}\perp P_{\theta_2}$. 为证此，设$\eta^{\theta_i}$是相应于$\theta_i$的参数系统，并按$P_{\theta_i}$考察该系统. 由假设，$\eta^{\theta_i}$是Gauss的，且关于$\xi$是线性的. 因此，$P_{\theta_1}\overset{\text{loc}}{\sim}P_{\theta_2}$. 按二断定理4.9.5，$P_{\theta_1}\sim P_{\theta_2}$与$P_{\theta_1}\perp P_{\theta_2}$两者必居其一. 假如$P_{\theta_1}\sim P_{\theta_2}$，则

$$P_{\theta_1}(\hat{\theta}_n\to\theta_1)=1\Rightarrow P_{\theta_2}(\hat{\theta}_n\to\theta_1)=1.$$

但 $P_{\theta_2}(\hat{\theta}_n\to\theta_2)=1$. 这就导致矛盾. 故

$$P_{\theta_1}\perp P_{\theta_2}(\theta_1\neq\theta_2).$$

系统η可以改写成如下形式：

$$\eta_{n+1}=\theta\eta_n+\sqrt{v_{n+1}}\varepsilon_{n+1},$$

其中$\varepsilon_n\sim N(0,1)$，ε为独立的r. v. 列. 现在，令

$\tilde{E}=E_{\theta_1}$，$E=E_{\theta_2}$，$\tilde{b}_n=\sqrt{v_{n+1}}$，$b_n=\sqrt{v_{n+1}}$，$\tilde{a}_n(x)=\theta_1x_n$，$a_n(x)=\theta_2x_n$，$\varphi_n(x)=(\theta_2-\theta_1)x_n$，$\tilde{E}\left(\dfrac{\varphi_n^2(x)}{\tilde{b}_n^2}\right)=(\theta_2-\theta_1)^2\dfrac{\tilde{E}x_n^2}{v_{n+1}}$，

则按定理4.9.5，有

$$P_{\theta_1} \perp P_{\theta_2} \Leftrightarrow \sum_n (\theta_2 - \theta_1)^2 \frac{\hat{E} x_n^2}{v_{n+1}} = \infty (\theta_1 \neq \theta_2).$$

现在，选取 $\theta_1=0, \theta_2 \neq 0$，则 $x_N^2=\xi_n^2 (\bmod P_{\theta_1})$. 注意

$$\widetilde{E}\, \xi_n^2 = E\xi_n^2 = v_n.$$

故　$P \perp P_{\theta_2} \Rightarrow \sum_n \frac{v_n}{v_{n+1}} = \infty. \ (\theta_2 \neq 0).$

第五章　最优停止规则

§5.1　最优停止问题(I)

设 $Z=(Z_n,n\geqslant 1)$是定义在概率空间$(\Omega,\mathscr{F},P)$上取值于状态空间$(E,\mathscr{E})$且适应于$(\mathscr{F}_.)=(\mathscr{F}_n,n\geqslant 1)$的随机序列. E-值 r. v. Z_n 可以解释为一个博弈者在第 n 轮博弈之后，其累计损失为 Z_n. 这就提出了一个问题：在什么情况下由停止博弈所造成的损失最小. 从概率论观点看，这种最小损失应指在平均意义下.

设 $\mathscr{M}=\{\tau:\tau$ 为关于$(P,\mathscr{F}_.)$的有限停时$\}$. 现在，要讨论的问题是：存不存在 $\tau^*\in\mathscr{M}$，使得

$$EZ_{\tau^*}\leqslant EZ_\tau\quad(\forall\,\tau\in\mathscr{M}).$$

即是说：能否找到一个有限停时 τ^*，使得博弈者相应的损失就其平均而言最小. 此问题并非总是有解，也不总是容易求解. 往往为求解需要将要求放宽到某种可接受的程度.

定义 5.1.1　假如对$\forall\,\tau\in\mathscr{M}$，$EZ_\tau$ 存在，且对$\forall\,\varepsilon>0$，存在 $\tilde{\tau}\in\mathscr{M}$，使得

$$EZ_{\tilde{\tau}}\leqslant EZ_\tau+\varepsilon\quad(\forall\,\tilde{\tau}\in\mathscr{M}),$$

则称 $\tilde{\tau}$ 为 Z 的 ε-最优停时. 简称 0-最优停止为最优停时.

定义 5.1.2　设 Z 为广义下鞅. 若对$\forall\,\tau\in\mathscr{M}$，$EZ_\tau$ 存在，且对$\forall\,n\geqslant 1$，有

$$E(Z_\tau/\mathscr{F}_n)\geqslant Z_n\ 于集(\tau\geqslant n)\ 上(\text{a.s}),$$

则称 Z 为正则广义下鞅.

按定理 4.6.5，若 $E(\sup\limits_n Z_n^+)<\infty$，则当 $Z\in GM_s(P,\mathscr{F}_.)$时，对$\forall\,n\geqslant 1$，有

$$E(Z_{n\vee\tau}/\mathscr{F}_n)\geqslant Z_n(\text{a.s})$$

注意，$(\tau\geqslant n)\in\mathscr{F}_n$. 于是，

$$1_{(\tau\geqslant n)}E(Z_\tau/\mathscr{F}_n)=E(1_{(\tau\geqslant n)}Z_{n\vee\tau}/\mathscr{F}_n)$$
$$\geqslant Z_n1_{(\tau\geqslant n)}(\text{a.s}).$$

引理 5.1.1 设 EZ_τ 存在($\forall\ \tau\in\mathscr{M}$)，且 $Z\in GM_s(P,\mathscr{F}.)$. 则 Z 正则的充要条件是

$$E(Z_\tau/\mathscr{F}_\sigma)\geqslant Z_\sigma(\text{a.s})\ \text{于集}(\sigma\leqslant\tau)\ \text{上}(\forall\ \sigma,\tau\in\mathscr{M}).$$

证 充分性显然. 下证必要性. 显然，

$$(\sigma\leqslant\tau)=\bigcup_n(\sigma=n)(n\leqslant\tau).$$

由 Z 的正则性知

$$1_{(\tau\geqslant n)}E(Z_\tau/\mathscr{F}_\sigma)1_{(\sigma=n)}=E(1_{(\tau\geqslant n)}Z_{n\vee\tau}/\mathscr{F}_n)1_{(\sigma=n)}$$
$$=1_{(\tau\geqslant n)}E(Z_{\tau\vee n}/\mathscr{F}_n)1_{(\sigma=n)}$$
$$\geqslant 1_{(\tau\geqslant n)}Z_n\cdot 1_{(\sigma=n)}=1_{(\tau\geqslant n)}Z_\sigma 1_{(\sigma=n)}$$
$$\Rightarrow E(Z_\tau/\mathscr{F}_\sigma)\geqslant Z_\sigma(\text{a.s})\ \text{于集}(\sigma\leqslant\tau)\ \text{上}.$$

注 1 此引理表明：广义下鞅的正则性等价于有限停时代换的不变性. 在第四章 §5 注 3 的例中，η 为非负鞅，且 $E\eta_n=1(\forall\ n\geqslant 1)$，$\eta_n\xrightarrow{\text{a.s}}0$. 在那里已证明：$\eta$ 不一致可积，且停时代换不能保证鞅性质不变. 但这个鞅还有下列特点：

(i) 对$\forall\ \varepsilon>0$，令 $\tau^*=\inf\{n\geqslant 1:\eta_n\leqslant\varepsilon\}$，$\inf\{\varnothing\}=\infty$，则 $\tau^*\in\mathscr{M}$；

(ii) $E\eta_{\tau^*}\leqslant\varepsilon$. 从而 $E\eta_{\tau^*}\leqslant E\eta_\tau+\varepsilon(\forall\ \tau\in\mathscr{M})$. 这表明：$\tau^*$ 为 η 的 ε-最优停时($\varepsilon>0$)；

(iii) 不存在有限最优停时. 事实上，若存在 $\tau'\in\mathscr{M}$，使得 $E\eta_{\tau'}\leqslant E\eta_\tau(\forall\ \tau\in\mathscr{M})$，则必有

$$E\eta_{\tau'}\leqslant E\eta_{\tau^*}\leqslant\varepsilon.\Rightarrow E\eta_{\tau'}=0.$$

然而，τ' 为有限停时，因此，$E\eta_{\tau'}>0$. 这就出现矛盾.

此例表明：当广义下鞅不是正则的时候，相应的最优停时问题在 $\mathscr{M}$ 范围内可以无解.

注 2 正则广义下鞅的最优停时为 $\tau'=1$. 这可由引理 5.1.1

立即得出. 也是停时代换不变性的一个必然结果. 具有有限个元素的广义下鞅的最优停时问题总有解. 基于这一点，为保证问题有解，可以放宽正则性要求，即$\exists$ 整数$k_0\geqslant 1$，使得

$$E(Z_\tau/\mathscr{F}_{k_0}) \geqslant Z_{k_0}(\text{a. s})\ \text{于}(\tau \geqslant k_0)\ \text{上}(\forall\ \tau \in \mathscr{M}).$$

这个条件比正则性条件要弱. 事实上，易证，对$\forall\ k\leqslant k_0$，有$E(Z_\tau/\mathscr{F}_k)\geqslant Z_k(\text{a. s})$于集$(\tau\geqslant k)$上. 然而，对$\forall\ k>k_0$，不能证明：正则性条件成立.

§5.2 最大广义下鞅

在上一节已知道：正则广义下鞅Z有最优停时$\tau'=1$. 而非正则广义下鞅Z在某些条件下，存在着ε-最优停时. 现在，要借助广义下鞅的特性来讨论一般过程的最优停时问题. 这里，最大广义下鞅将起着特别重要的作用. 符号$GA_s(\mathscr{F}.)$表示广义适应随机序列全体.

定义 5.2.1 设$Z=(Z_n,n\geqslant 1)\in GA_s(\mathscr{F}.)$，$\xi=(\xi_n,n\geqslant 1)\in GA_s(\mathscr{F}.)$. 若

(i) $\xi\in GM_s(P,\mathscr{F}.)$；

(ii) $\xi<Z$，即$\xi_n\leqslant Z_n(\text{a. s})(n\geqslant 1)$；

(iii) $\eta<\xi$ ($\forall\ \eta\in GM_s(P,\mathscr{F}.)$，$\eta<Z$)，

则称ξ为关于Z的最大广义下鞅.

条件A：$\exists$ r. v. $u\leqslant 0$，$\dashv Eu>-\infty$，且$Z_n\geqslant u(\text{a. s})(\forall\ n\geqslant 1)$.

注 1 若Z满足条件A，则说Z受u下控. 条件A的作用在于保证相应于Z的正则广义下鞅类是非空的. 这是因为一致可积鞅必正则. 显然，条件A等价于：$E(\inf\limits_n Z_n)>-\infty$或$E(\sup\limits_n Z_n^-)<\infty$.

定理 5.2.1 若Z满足条件A，则相对于Z的正则广义下鞅存在.

证 设$\mathscr{U}$为Z的正则广义下鞅全体. 按注 1，$\mathscr{U}$是非空的.

令

$$\xi = \text{ess sup}\{\eta : \eta \in \mathscr{U}\},$$

则 $\xi \in \mathscr{U}$. 从而，定理成立.

事实上，对$\forall\ \sigma, \tau \in \mathscr{M}$，且 $P(\sigma \leqslant \tau) = 1$. 按正则性，有

$$E(\eta_\tau / \mathscr{F}_\sigma) \geqslant \eta_\sigma (\text{a.s}) \quad (\forall\ \eta \in \mathscr{U})$$

$$\Rightarrow E(\xi_\tau / \mathscr{F}_\sigma) \geqslant \eta_\sigma (\text{a.s}) \quad (\forall\ \eta \in \mathscr{U})$$

$$\Rightarrow E(\xi_\tau / \mathscr{F}_\sigma) \geqslant \text{ess sup}\{\eta_\sigma : \eta \in \mathscr{U}\} = \xi_\sigma (\text{a.s}).$$

即 $\xi \in \mathscr{U}$. 显然，$\xi < Z$. 故 ξ 为关于 Z 的最大正则广义下鞅.

注 2 类似的方法也可以作出关于 Z 的最大广义下鞅(不必正则).

定理 5.2.2 设 $Z = (Z_n, n \geqslant 1) \in GA_s(\mathscr{F}.)$满足条件 A. ξ 为关于 Z 的最大正则广义下鞅，则 ξ 具有下列性质：

(i) $\xi_n = Z_n \wedge E(\xi_{n+1} / \mathscr{F}_n)$ (a.s)$(n \geqslant 1)$;

(ii) $\varliminf\limits_{n\to\infty} \xi_n = \varliminf\limits_{n\to\infty} Z_n$(a.s).

证 首先证(i). 按条件，有

$$P(\xi_r \leqslant Z_r \wedge E(\xi_{r+1}) / \mathscr{F}_r)) = 1.$$

假如(i)不真，则$\exists\ r \geqslant 1$，$\ni$

$$P(\xi_r < Z_r \wedge E(\xi_{r+1} / \mathscr{F}_r)) > 0. \tag{5.2-1}$$

记 $I_B(n) = \begin{cases} 1, n \in B, \\ 0, n \notin B. \end{cases}$ $B_1 = \{k : k > r\}, B_2 = \{r\}, B_3 = \{k : k < r\}$.

$$\xi'_n = \xi_n I_{B_1}(n) + Z_r \wedge E(\xi_{r+1} / \mathscr{F}_r) I_{B_2}(n) + E(u / \mathscr{F}_n) I_{B_3}(n).$$

则$E(\xi'_{n+1} / \mathscr{F}_n)$

$$\begin{aligned} &= E(\xi_{n+1} I_{B_1}(n+1) \\ &\quad + Z_r \wedge E(\xi_{r+1} / \mathscr{F}_r) I_{B_2}(n+1) \\ &\quad + E(u / \mathscr{F}_{n+1}) I_{B_3}(n+1) / \mathscr{F}_n) \\ &\geqslant \xi_n I_{B_1}(n) + E(\xi_{n+1} / \mathscr{F}_n) I_{B_2}(n) \\ &\quad + E(u / \mathscr{F}_n) I_{B_2}(n+1) + E(u / \mathscr{F}_n) I_{B_3}(n+1) \\ &\geqslant \xi_n I_{B_1}(n) + Z_r \wedge E(\xi_{r+1} / \mathscr{F}_r) I_{B_2}(n) \\ &\quad + E(u / \mathscr{F}_n) I_{B_3}(n) = \xi'_n. \quad (\forall\ n \geqslant 1). \end{aligned}$$

即 ξ' 为广义下鞅. 为验证 ξ' 的正则性. 按本章 § 1 注 2,仅需验证 $n>r$ 的情形. 设 $\tau\in\mathscr{M}$,则

$$1_{(\tau\geqslant n)}E(\xi'_\tau/\mathscr{F}_n)=E(\xi'_\tau 1_{(\tau\geqslant n)}/\mathscr{F}_n)=E(\xi_\tau 1_{(\tau\geqslant n)}/\mathscr{F}_n)$$
$$=1_{(\tau\geqslant n)}E(\xi_\tau/\mathscr{F}_n)\geqslant 1_{(\tau\geqslant n)}\xi_n=1_{(\tau\geqslant n)}\xi'_n(\text{a. s}).$$

故 ξ' 正则. 从而,$\xi'\in\mathscr{U}$. 然而,$\xi'<\xi$. 于是,

$$Z_r\wedge E(\xi_{r+1}/\mathscr{F}_r)=\xi'_r\leqslant\xi_r(\text{a. s}).$$

故(5. 2-1)式与此式矛盾,即(5. 2-1)式不成立.

下证(ii). 为此,令

$$\zeta_n=E(\inf_{k\geqslant n}Z_k/\mathscr{F}_n).$$

则
$$E(\zeta_{n+1}/\mathscr{F}_n)=E(\inf_{k\geqslant n+1}Z_k/\mathscr{F}_n)$$
$$\geqslant E(\inf_{k\geqslant n}Z_k/\mathscr{F}_n)=\zeta_n(\text{a. s}).$$

即 ζ 为广义下鞅. 令

$$\zeta_n^c=E((\inf_{k\geqslant n}Z_k)\wedge c/\mathscr{F}_n)\ (c>0).$$

则 ζ^c 为一致可积下鞅,且受两个一致可积鞅的控制. 按推论 4. 3. 2 和定理 4. 3. 1,有

$$E(\zeta_\tau^c/\mathscr{F}_\sigma)\geqslant\zeta_\sigma^c(\text{a. s})\ (\forall\ \sigma,\tau\in\mathscr{M},\text{且}\ \sigma\leqslant\tau\ (\text{a. s})).$$

按单调收敛定理,得

$$E(\zeta_\tau/\mathscr{F}_\sigma)=\lim_{c\to\infty}E(\zeta_\tau^c/\mathscr{F}_\sigma)\geqslant\lim_{c\to\infty}\zeta_\sigma^c=\zeta_\sigma(\text{a. s}).$$

这表明:ζ 是正则的,即 $\zeta\in\mathscr{U}$. 按定理 4. 6. 3(ii),并注意:$\zeta<\xi$,有

$$\varliminf_{n\to\infty}\xi_n\geqslant\varliminf_{n\to\infty}\zeta_n\geqslant\lim_{n\to\infty}E(\inf_{k\geqslant m}Z_k/\mathscr{F}_n)$$
$$\geqslant\inf_{k\geqslant m}Z_k(\text{a. s})\ (\forall\ m\geqslant 1)$$
$$\Rightarrow\varliminf_{n\to\infty}\xi_n\geqslant\varlimsup_{n\to\infty}Z_n(\text{a. s}).\Rightarrow\text{结论(ii) 成立}.$$

定理 5. 2. 3 设 $Z=(Z_n,n\geqslant 1)\in GA_s(\mathscr{F}.)$满足条件 A. ξ 为关于 Z 的最大正则广义下鞅. 令

$$\tau^*=\inf\{n\geqslant 1:\xi_n\geqslant Z_n-\varepsilon\},\inf\{\varnothing\}=\infty.$$

则 τ^* 是 Z 的 ε-最优停时($\varepsilon>0$). 此外

$$E\xi_n=\inf\{EZ_\tau:\tau\in\mathscr{M},\text{且}\ P(\tau\geqslant n)=1\}(\forall\ n\geqslant 1).$$

证 先证 $\tau^*\in\mathscr{M}$. 为此,令 $\eta_n=\xi_{n\wedge\tau^*}(n\geqslant 1)$,则 η 为广义鞅.

事实上,

$\eta_n=\eta_{n+1}$于$(\tau^*\leqslant n)$上,

$$\int_{A\cdot(\tau^*\leqslant n)}\eta_n\mathrm{d}P=\int_{A\cdot(\tau^*\leqslant n)}\eta_{n+1}\mathrm{d}P(\forall\ A\in\mathscr{F}_n).$$

注意,在$(\tau^*>n)$上,有$\xi_n<Z_n$. 按定理 5.2.2,有

$$\xi_n=Z_n\wedge E(\xi_{n+1}/\mathscr{F}_n)$$

$$\Rightarrow\xi_n=E(\xi_{n+1}/\mathscr{F}_n)\ (\text{a.s})\text{ 于}(\tau^*>n)\text{ 上}$$

$$\Rightarrow\int_{A\cdot(\tau^*>n)}\xi_n\mathrm{d}P=\int_{A\cdot(\tau^*>n)}\xi_{n+1}\mathrm{d}P\ (\forall\ A\in\mathscr{F}_n)$$

$$\Rightarrow\int_A\eta_n\mathrm{d}P=\int_A\eta_{n+1}\mathrm{d}P\ (\forall\ A\in\mathscr{F}_n),$$即η为广义鞅.

设u为条件A中的 r.v. 定义$\mu_{\tau^*}=E(u/\mathscr{F}_{\tau^*})$,则 $\mu_{\tau^*}\leqslant 0$,
$E\mu_{\tau^*}=Eu>-\infty$,且

$$E(\mu_{\tau^*}/\mathscr{F}_n)\leqslant\eta_n(\text{a.s})\ (\forall\ n\geqslant 1).$$

最后的这个不等式证明如下:由ξ的最大性知:

$$\xi_n\geqslant E(u/\mathscr{F}_n)\ (\text{a.s})\ (n\geqslant 1)$$

$$\Rightarrow\xi_{n\wedge\tau^*}\geqslant E(u/\mathscr{F}_{n\wedge\tau^*})$$

$$\Rightarrow\eta_n=\lim_{m\to\infty}E(\eta_{n+m}/\mathscr{F}_n)$$

$$\geqslant\lim_{m\to\infty}E(E(\mu_{\tau^*}/\mathscr{F}_{(n+m)\wedge\tau^*})/\mathscr{F}_n)$$

$$=E(\mu_{\tau^*}/\mathscr{F}_n).$$

按定理 4.6.4,$\eta_\infty=\lim\limits_{n\to\infty}\eta_n$(a.s)存在,$\eta_\infty$在集$\{\eta_1<\infty\}$上有限,且
$E(\eta_\infty/\mathscr{F}_1)\leqslant\eta_1$(a.s). 按$\tau^*$之定义,

$$(\tau^*=\infty)\subseteq\{\xi_1<\infty\}=\{\eta_1<\infty\}.$$

于是,η_∞于集$(\tau^*=\infty)$上(a.s)有限. 注意

$$\lim_{n\to\infty}\xi_n=\lim_{n\to\infty}\eta_n\text{ 于}(\tau^*=\infty)\text{ 上}.$$

故$\lim\limits_{n\to\infty}\zeta_n=\lim\limits_{n\to\infty}\xi_n$于$(\tau^*=\infty)$上(a.s)有限. 按定理 5.2.2,有

$$P(\tau^*=\infty)=P(\{\underline{\lim}\xi_n\leqslant\underline{\lim}Z_n-\varepsilon<\underline{\lim}Z_n\}\cdot(\tau^*=\infty))$$

$$\leqslant P(\underline{\lim}\xi_n<\underline{\lim}Z_n)=0.$$

故 τ^* 为有限停时，即 $\tau^* \in \mathscr{M}$.

现在证明：τ^* 为 Z 的 ε-最优停时($\varepsilon>0$). 事实上，

$$\tau^* \in \mathscr{M} \Rightarrow \xi_{\tau^*} = \eta_\infty (\text{a.s}).$$

按 ξ 的正则性，有 $E(\xi_{\tau^*}/\mathscr{F}_1) \geqslant \xi_1 (\text{a.s})$. 另一方面，

$$E(\xi_{\tau^*}/\mathscr{F}_1) = E(\eta_\infty/\mathscr{F}_1) \leqslant \eta_1 = \xi_1 (\text{a.s})$$

$$\Rightarrow E(\xi_{\tau^*}/\mathscr{F}_1) = \xi_1 (\text{a.s}).$$

按 τ^* 之定义，有

$$E(Z_{\tau^*}/\mathscr{F}_1) \leqslant E(\xi_{\tau^*}/\mathscr{F}_1) + \varepsilon = \xi_1 + \varepsilon$$

$$\Rightarrow EZ_\tau \geqslant E\xi_\tau \geqslant E\xi_1 = E\xi_{\tau^*} \geqslant EZ_{\tau^*} - \varepsilon. \ (\forall \tau \in \mathscr{M}).$$

故 τ^* 为 Z 的 ε-最优停时.

最后讨论 $E\xi_n$. 注意

$$EZ_\tau \geqslant E\xi_1 \geqslant EZ_{\tau^*} - \varepsilon \ (\forall \tau \in \mathscr{M})$$

$$\Rightarrow E\xi_1 = \inf\{EZ_\tau : \tau \in \mathscr{M}\} = \inf\{EZ_\tau : \tau \in \mathscr{M}, P(\tau \geqslant 1) = 1\}.$$

若将起始项 Z_1 改成 Z_n，则类似地作法可得

$$E\xi_n = \inf\{EZ_\tau : \tau \in \mathscr{M}, P(\tau \geqslant n) = 1\}.$$

注 3 满足条件 A 的 Z 所相应的最大正则广义下鞅 ξ 具有如下性质：$\exists\ \tau^* \in \mathscr{M}$，使得

$\eta = (\eta_n, n \geqslant 1)$ 为广义鞅，$\eta_n = \xi_{n \wedge \tau^*}$.

注 4 满足条件 A 的 Z 之 ε-最优停时问题可以通过找相应的最大正则广义下鞅 ξ 来解决. 有两个途径，其一是通过关于 Z 的正则广义下鞅类求其本质上确界. 这个处理过程一般比较复杂. 其二是通过定理 5.2.2 中的方程构作 ξ，即解如下问题：

$$\begin{cases} \xi_n = Z_n \wedge E(\xi_{n+1}/\mathscr{F}_n) \ (\forall n \geqslant 1), \\ \underline{\lim}\,\xi_n = \underline{\lim}\,Z_n (\text{a.s}). \end{cases} \tag{5.2-2}$$

由此提出下列问题：

(i) 方程(5.2-2)的解是否唯一？

(ii) 方程(5.2-2)的解是否正则？

(iii) 如何由方程(5.2-2)找出 Z 的最大正则广义下鞅解？

令 $\mu_n=E(u/\mathscr{F}_n)$，则 $\mu_n\leqslant Z_n$(a. s)($\forall\ n\geqslant 1$). 因 $0\geqslant u\in L'$，而知 μ 为正则鞅. 并显然满足(5. 2-2)中的方程. 但一般不满足(5. 2-2)中的条件. 以后会讨论定理 5. 2. 2 的逆问题.

注 5 假定 $Z\in GA_s(\mathscr{F}.)$ 为一个可观测过程，每一步之后的输出为 $g_n=f(z_1,z_2,\cdots,z_n)$（累计输出）或 $f(z_n)$（当步输出）. 如果将 g_n 当作某种损失来考虑，那么我们希望找到一个停止规则 $\tau^*\in\mathscr{M}$，使得：

$$Eg_{\tau^*}\leqslant Eg_\tau+\varepsilon.\quad(\forall\ \tau\in\mathscr{M}).$$

本节在解决关于 Z 的相应问题之过程中，仅要求 Z 为适应序列. 然而，g 也是适应序列，因此，用 g 代替 Z，本节的构造方法仍适用. 如果观测过程 Z 具有 Markov 性，则最优停时问题在结构上必然具有某些新的特性. Markov 过程论告诉我们：在适当条件下，通过增维处理可将非齐次 Markov 过程转化成齐次 Markov 过程. 基于这一点. 以后侧重详细地介绍与 Markov 过程有关的最优停时问题.

§5. 3 最优停止问题(Ⅱ)

设 $\xi=(\xi_n,\mathscr{F}_n,n\in N)$($N=\{0,1,2,\cdots,\}$)为定义在概率空间$(\Omega,\mathscr{F},P_x)$上取值于$(E,\mathscr{E})$的齐次 Markov 链($x\in E$)，这里假定 σ-代数 $\mathscr{E}$ 含 E 中每个单点集，而

$$P_x(\xi_0\in B)=\begin{cases}1, & x\in B,\\ 0, & x\notin B.\end{cases}\ (\forall\ B\in\mathscr{E}),$$

即是说：此齐次 Markov 链的初始分布集中在点 $x\in E$. 本节讨论观测过程为齐次 Markov 链的最优停时问题. 关于 Markov 链的知识有大量的文献可供参考. 本书将在最后提供一个够用的简介以方便阅读.

前面讨论相当一般的最小损失问题. 本节以后将讨论最大效益问题. 但本质没有太大的变化，只是广义下鞅被代之以广义上

鞅. 讨论的对象具有 Markov 性. 定义下列常用符号:

$B=\{g:g=g(x),x\in E$,取值于$(-\infty,\infty]$,且$\mathscr{E}$-可测$\}$.

$\mathscr{M}^n=\{\tau:\tau$ 为停时,且 $\tau\leqslant n\}(n\in\mathrm{N})$.

$\mathscr{M}=\bigcup\limits_{n=0}^{\infty}\mathscr{M}^n$(有限停时类),

$\overline{\mathscr{M}}$ 为一般停时类.

$g(\xi_n)$表示时刻 n 的当步输出. 常称 g 为目标函数.

$$S_n(x)=\sup\{E_xg(\xi_\tau):\tau\in\mathscr{M}^n\}.$$

$$S(x)=\lim_{n\to\infty}S_n(x).$$

$$\overline{S}(x)=\sup\{E_xg(\xi_\tau):\tau\in\overline{\mathscr{M}}\}$$

$$\mathscr{M}_g^n=\{\tau\in\mathscr{M}^n:E_xg^-(\xi_\tau)<\infty(\forall\ x\in E)\}\ (g\in B).$$

$$\mathscr{M}_g=\{\tau\in\mathscr{M}:E_xg^-(\xi_\tau)<\infty(\forall\ x\in E)\}(g\in B).$$

$$g(\xi_\infty)=\overline{\lim_{n\to\infty}}g(\xi_n).$$

以后称函数 $S_n,S,\overline{S}$ 为 g 的报酬函数.

定义 5.3.1 若 $\tau_n^*\in\mathscr{M}^n$,且

$$S_n(x)=E_xg(\xi_{\tau_n^*})\ (\forall\ x\in E),$$

则称 τ_n^* 为类 $\mathscr{M}^n$ 中的最优停时(关于 g).

引理 5.3.1 对$\forall\ g\in B$,及 $n\in N$,类 $\mathscr{M}_g^n$ 和 $\mathscr{M}_g$ 均是非空的.

证 让 $\tau\equiv0$. 则 $\tau\in\mathscr{M}^n\subset\mathscr{M}$,且

$$E_xg(\xi_\tau)=E_xg(x)=g(x)>-\infty\ (\forall\ g\in B).$$

$$\Rightarrow\tau\equiv0\in\mathscr{M}_g^n\subset\mathscr{M}_g.$$

引理 5.3.2 $S_n(x)=\sup\{E_xg(\xi_\tau):\tau\in\mathscr{M}_g^n\}(g\in B,x\in E)$,即在 $S_n(x)$的定义中将 $\mathscr{M}^n$ 改成 $\mathscr{M}_g^n$,其值不变.

证 显然,有

$$S_n(x)\geqslant S'_n(x)=\sup\{E_xg(\xi_\tau):\tau\in\mathscr{M}_g^n\}.$$

因此,此引理等价于:$S_n(x)\leqslant S'_n(x)\ (\forall\ x\in E)$.

设 $x_0\in E,g\in B$,且对$\forall\ \tau\in\mathscr{M}^n$,$E_{x_0}g(\xi_\tau)$存在. 则此引理可由

如下命题推出：对$\forall\ \tau\in\mathscr{M}^n$，$\exists\ \sigma\in\mathscr{M}^n$，使得

(i) $E_{x_0}g(\xi_\sigma)\geqslant E_{x_0}g(\xi_\tau)$.

(ii) $E_xg^-(\xi_\sigma)<\infty\ (\forall\ x\in E)$.

事实上，(ii)成立表明：$\sigma\in\mathscr{M}_g^n$. 而(i)成立表明：

$$
\begin{aligned}
S_n(x_0) &= \sup\{E_{x_0}g(\xi_\tau);\tau\in\mathscr{M}^n\} \\
&\leqslant \sup\{E_{x_0}g(\xi_\sigma);\sigma\in\mathscr{M}_g^n\} = S'_n(x_0).
\end{aligned}
$$

设$\tau\in\mathscr{M}^n$. 若$E_{x_0}g^-(\xi_\tau)=\infty$，则按引理5.3.1，取$\sigma\equiv 0$，必使(i)和(ii)满足. 现假定：$E_{x_0}g^-(\xi_\tau)<\infty$. 定义

$$
\sigma(\omega)=\tau(\omega)\cdot 1_{\{\xi_0=x_0\}}.
$$

显然，由$\{\xi_0=x_0\}\in\mathscr{F}_0$知：$\sigma$为停时，且$\sigma\in\mathscr{M}^n$. 下面验证$\sigma$满足要求(i)和(ii).

$$
\begin{aligned}
E_{x_0}g(\xi_\tau) &= E_{x_0}[E_{x_0}(g(\xi_\tau)/\mathscr{F}_0)(1_{\{\xi_0=x_0\}}+1_{\{\xi_0\neq x_0\}})] \\
&= E_{x_0}(g(\xi_\tau)1_{\{\xi_0=x_0\}}) \\
&= E_{x_0}(g(\xi_\sigma)1_{\{\xi_0=x_0\}}) = E_{x_0}g(\xi_\sigma).
\end{aligned}
$$

$$
\begin{aligned}
E_xg^-(\xi_\sigma) &= E_x(g^-(\xi_\sigma)1_{\{\xi_0=x_0\}}) + E_x(g^-(\xi_\sigma)\cdot 1_{\{\xi_0\neq x_0\}}) \\
&= \left.\begin{cases} E_{x_0}g^-(\xi_\tau), & \text{当 } x=x_0 \text{ 时} \\ g^-(x), & \text{当 } x\neq x_0 \text{ 时}\end{cases}\right\}<\infty(\forall\ g\in B, x\in E),
\end{aligned}
$$

故σ满足(i)和(ii).

注1 因$\tau=0\in\mathscr{M}_g^n\subseteq\mathscr{M}_g$，因此，按上述引理，当$E_xg^\pm(\xi_\tau)=\infty$时，规定$E_xg(\xi_\tau)=-\infty$不会影响$S_n(x)$和$S(x)$的值. 以后我们不再说明：$E_xg(\xi_\tau)$是否有定义.

定义 5.3.2 如果对$\varepsilon\geqslant 0$，$\exists\ \tau_\varepsilon\in\mathscr{M}_g$，使得

$$
S(x)-\varepsilon\leqslant E_xg(\xi_{\tau_\varepsilon})\ (\forall\ x\in E)
$$

或者$E_xg(\xi_\tau)-\varepsilon\leqslant E_xg(\xi_{\tau_\varepsilon})\ (\forall\ x\in E,\tau\in\mathscr{M}_g)$，
则称τ_ε为g的一个(ε,S)-最优停时或ε-最优停时. 而称τ_0(即$\varepsilon=0$)为g的$(0,S)$-最优停时或最优停时.

定义 5.3.3 若$\exists\ \tau_\varepsilon\in\overline{\mathscr{M}}_g(\varepsilon\geqslant 0)$，使得

$$\overline{S}(x)-\varepsilon\leqslant E_x g(\xi_{\tau_\varepsilon})\ (\forall\ x\in E),$$

则称 τ_ε 为 g 的 $(\varepsilon,\overline{S})$-最优停时.

注 2 当 $\varepsilon>0$ 时, (ε,S)-最优停时亦为 $(\varepsilon,\overline{S})$-最优停时. 但反之一般不真. 另一方面, 当 $\varepsilon=0$ 时, $(0,S)$-最优停时可以不存在. 但 $(0,\overline{S})$-最优停时存在. 这可从 § 1 的例中看出.

§ 5.4 在类 $\mathscr{M}^n$ 中的最优停时

本节讨论观测过程为有限齐次 Markov 链 ξ 的最优停止问题. 定义下列符号:

$$B_1=\{g\in B: E_x g^-(\xi_1)<\infty\ (\forall\ x\in E)\}.$$

$Q: B_1\to R$ 之映射: $Qg(x)=g(x)\vee Tg(x)\ (\forall\ x\in E, g\in B_1)$.

$T: B_1\to R$ 之映射: $Tg(x)=E_x g(\xi_1)\ (\forall\ x\in E, g\in B_1)$.

$$Q^0 g(x)=g(x);$$

$$Q^n g(x)=Q(Q^{n-1}g(x))\ (\forall\ n\geqslant 1).$$

引理 5.4.1 $Q^n g^{(x)}=g(x)\vee TQ^{n-1}g(x)\ (\forall\ x\in E, g\in B_1, n\geqslant 1)$.

证 显然

$$T(g\vee Tg)(x)\geqslant Tg(x), \text{即 } TQg(x)\geqslant Tg(x)$$

$$\Rightarrow Q^2 g(x)=Qg(x)\vee TQg(x)$$

$$=(g(x)\vee Tg(x))\vee TQg(x)=g(x)\vee TQg(x).$$

采用归纳法立即得所要的结论.

引理 5.4.2

$$E_x g(\xi_\tau)\leqslant Q^n g(x)\ (\forall\ x\in E, g\in B, \tau\in\mathscr{M}^n, n\in\mathrm{N})$$

证 按上节注 1, 若 $E_x g(\xi_\tau)$ 不存在, 就定义它为 $-\infty$. 在这种情形下, 引理自然成立. 因此, 可以假定 $\tau\in\mathscr{M}^n_g$. 若 $n=0$, 则由 $\tau\leqslant n$ 知: $\tau=0$. 从而, 引理成立. 现在考虑 $n>0$. 记 $A=(\tau=n)$, 则

$$A=\Omega\setminus(\tau<n)\in\mathscr{F}_{n-1}.$$

利用这一事实及 ξ 的齐次 Markov 性，得

$$E_x g(\xi_\tau) = E_x g(\xi_\tau)(1_A + 1_{A^C})$$
$$= E_x(g(\xi_{\tau\wedge(n-1)})1_{A^C} + g(\xi_n)1_A);$$
$$E_x(g(\xi_n)1_A) = E_x(E(g(\xi_n)1_A/\mathscr{F}_{n-1}))$$
$$= E_x(1_A E_{\xi_{\tau\wedge(n-1)}} g(\xi_1))$$
$$\Rightarrow E_x g(\xi_\tau) \leqslant E_x(g(\xi_{\tau\wedge(n-1)}) \vee Tg(\xi_{\tau\wedge(n-1)}))$$
$$= E_x Qg(\xi_{\tau\wedge(n-1)}).$$

令 $\sigma=\tau\wedge(n-1)$，$g_1(x)=Qg(x)$，则 $\sigma\in\mathscr{M}_g^{n-1}$，$g_1\in B$. 同理可得

$$E_x g_1(\xi_\sigma) \leqslant E_x Qg_1(\xi_{\sigma\wedge(n-2)}) = E_x Q^2 g(\xi_{\tau\wedge(n-2)})$$
$$\Rightarrow E_x g(\xi_\tau) \leqslant E_x g_1(\xi_\sigma) \leqslant E_x Q^2 g(\xi_{\tau\wedge(n-2)})$$
$$\Rightarrow E_x g(\xi_\tau) \leqslant E_x Q^n g(\xi_0) = Q^n g(x).$$

引理 5.4.3 设 $g\in B_1$，有

$$\sigma_n = \inf\{0\leqslant k\leqslant n: Q^{n-k}g(\xi_k) = g(\xi_k)\}, \inf\{\varnothing\} = \infty.$$

则(i) $\sigma_n\in\mathscr{M}_g^n (\forall\, n\in N)$；

(ii) $Q^n g(x)=E_x g(\xi_{\sigma_n})\ (\forall\, x\in E).\ (\forall\, n\in \mathrm{N})$.

证 σ_n 为停时是显然的，又

$$\{Q^0 g(\xi_n) = g(\xi_n)\} = \Omega. \Rightarrow \{\sigma_n \leqslant n\} = \Omega \Rightarrow \sigma_n \in \mathscr{M}^n.$$
$$g \in B_1 \Rightarrow Tg(x) > -\infty,\ g(x) > -\infty$$
$$(\forall\, x \in E). \Rightarrow Qg(x) > -\infty\ (\forall\, x \in E).$$

按引理 5.4.1，有

$$Q^n g(x) \geqslant Tg(x) > -\infty\ (\forall\, x \in E).$$

若结论(ii)成立，则有 $E_x g(\xi_{\sigma_n}) > -\infty\ (\forall\, x\in E)$. 从而，$\sigma_n\in\mathscr{M}_g^n$. 剩下的问题就是证明(ii)成立.

采用归纳法，当 $n=0$ 时，(ii)显然成立. 假定(ii)对 $n-1$ 成立. 下面证它对 n 亦成立. 为此，定义平移算子 θ：$\theta(f(\xi_k,\cdots,\xi_m))=f(\xi_{k+1},\cdots,\xi_{m+1})\ (\forall\, m\geqslant k)$. 则

$$\sigma_n = \theta(\sigma_{n-1}) \cdot 1_{(\sigma_n>0)}.$$

利用这一事实及 Markov 性，可得

$$\begin{aligned}E_x g(\xi_{\sigma_n}) &= E_x(g(\xi_0)1_{(\sigma_n=0)}) + E_x(g(\xi_{\theta(\sigma_{n-1})})\cdot 1_{(\sigma_n>0)}) \\ &= E_x(g(\xi_0)1_{(\sigma_n=0)}) + E_x(1_{(\sigma_n>0)}E_{\xi_1}g(\xi_{\sigma_{n-1}})) \\ &= E_x(Q^n g(\xi_0)1_{(\sigma_n=0)}) + E_x(1_{(\sigma_n>0)}E_{\xi_0}Q^{n-1}g(\xi_1)).\end{aligned}$$

按 σ_n 之定义，$g(\xi_0)<Q^n g(\xi_0)$ 于 $(\sigma_n>0)$ 上. 按引理 5.4.1，有

$$Q^n g(\xi_0) = g(\xi_0) \vee TQ^{n-1}g(\xi_0) = TQ^{n-1}g(\xi_0) \text{ 于 } (\sigma_n > 0) \text{ 上}.$$

综上分析，最后得

$$\begin{aligned}E_x g(\xi_{\sigma_n}) &= E_x(Q^n g(\xi_0)1_{(\sigma_n=0)} + 1_{(\sigma_n>0)}Q^n g(\xi_0)) \\ &= E_x Q^n g(\xi_0) = Q^n g(x).\end{aligned}$$

定理 5.4.1 设 $g\in B_1$，则对 $\forall\, n\in \mathrm{N}$ 及 $x\in E$，有

(i) $S_n(x)=Q^n g(x)$；

(ii) $S_n(x)=g(x)\vee TS_{n-1}(x)$；

(iii) $S_n(x)=E_x g(\xi_{\tau_n^*})$，其中

$$\tau_n^* = \min\{0\leqslant k\leqslant n : S_{n-k}(\xi_k) = g(\xi_k)\},$$

$\tau_n^*\in \mathscr{M}_g^n$，且为最优停时；

(iv) $\mathscr{M}_g^m=\mathscr{M}^m\,(0\leqslant m\leqslant n)$ 当 $E_x g^-(\xi_k)<\infty\ (0\leqslant k\leqslant n, x\in E)$ 时.

证 按引理 5.4.2 和 5.4.3，得

$$\begin{aligned}Q^n g(x) &= E_x g(\xi_{\sigma_n}) \\ &= \sup\{E_x g(\xi_\tau) : \tau\in\mathscr{M}^n\} = S_n(x)\Rightarrow(\mathrm{i}).\end{aligned}$$

按引理 5.4.1 及结论(i)，立即可得结论(ii).

按引理 5.4.3 及结论(i)，立即可得结论(iii).

剩下结论(iv)需要证明. 为此，设 $\tau\in\mathscr{M}^n$，则

$$\begin{aligned}E_x g^-(\xi_\tau) &= \sum_{k=0}^n E_x(g^-(\xi_k)1_{(\tau=k)}) \\ &\leqslant \sum_{k=0}^n E_x g^-(\xi_k) < \infty\end{aligned}$$

$$\Rightarrow \tau\in\mathscr{M}_g^n, \text{即 } \mathscr{M}^n\subset\mathscr{M}_g^n. \Rightarrow(\mathrm{iv}).$$

注 1 令 $\Gamma_k^n=\{x\in E : S_{n-k}(x)=g(x)\}\,(0\leqslant k\leqslant n)$，则在定理

5.4.1 中的最优停时 τ_n^* 可借助于 Γ 表示如下：

$$\tau_n^* = \min\{0 \leqslant k \leqslant n : \xi_k \in \Gamma_k^n\}.$$

$\Gamma_n^* = E$ 表明：观测过程在时刻 n 必停止. 而 Γ_k^* 表示观测过程的停止域. 自然，$C_k^n = E \backslash \Gamma_k^*$ 表示观测过程的继续域，按引理 5.4.1，继续观测域 C 具有如下结构：

$$C_n^n = \varnothing; C_{n-1}^n = \{x \in E : S_1(x) > g(x)\}; \cdots$$

$$C_0^n = \{x \in E : S_n(x) > g(x)\}.$$

从此表示中可以看出报酬函数 S_k 随 k 增加而增加. 现在，设 $\Gamma^n = (\Gamma_k^n, 0 \leqslant k \leqslant n) \subset \mathscr{E}$，即每个 Γ_k^n 为 $\mathscr{E}$ 中的可测子集，定义停时：

$$\tau_{\Gamma^n} = \min\{0 \leqslant k \leqslant n : \xi_k \in \Gamma_k^n\}, \min\{\varnothing\} = n.$$

则 $\tau_{\Gamma^n} \in \mathscr{M}^n$；令

$$\mathscr{S}^n = \{\tau \in \mathscr{M}^n : \tau = \tau_{\Gamma^n}, \text{其中 } \Gamma^n \subset \mathscr{E}\},$$

则
$$S_n(x) = \sup\{E_x g(\xi_\tau) : \tau \in \mathscr{S}^n\}.$$

这就给出了最优停时问题的另一种表达形式.

引理 5.4.4 设 $g \in B_1$，且 $E_x g^-(\xi_k) < \infty \ (0 \leqslant k \leqslant n)$，则

$$(S_{n-k}(\xi_k), 0 \leqslant k \leqslant n) \in GM^s(P_x, \mathscr{F}_.) \ (\forall\ x \in E),$$

即 $(S_{n-k}(\xi_k), 0 \leqslant k \leqslant n)$ 关于 P_x 为广义上鞅 $(\forall\ x \in E)$.

证 按定理 5.4.1(iii)～(iv)，有

$$S_m(x) = E_x g(\xi_{\tau_m^*}) > -\infty \ (0 \leqslant m \leqslant n).$$

剩下的问题在于验证上鞅性. 这可从下述关系中得到证实：

$$E_x(S_{n-(k+1)}(\xi_{k+1}) / \mathscr{F}_k) = E_{\xi_k} S_{n-k-1}(\xi_1)$$

$$= TS_{n-k-1}(\xi_k) \leqslant S_{n-k}(\xi_k) (\mathrm{mod} P_x) (\forall\ x \in E)$$

(按定理 5.4.1(ii)).

注 2 如果 $\eta^* = (\eta_k, 0 \leqslant k \leqslant n) \in GM^s(P_x, \mathscr{F}_.)$，且 $g(\xi_k) \leqslant \eta_k$ $(\mathrm{mod} P_x)(0 \leqslant k \leqslant n)$. 那么是否有

$$S_{n-k}(\xi_k) \leqslant \eta_k (\mathrm{mod} P_x)(0 \leqslant k \leqslant n)?$$

这个问题将在下一节讨论.

下面介绍一种所谓"向后递归原理". 首先定义下列符号.

$$\mathscr{M}^{m,n} = \{\tau \in \mathscr{M}^n : (m \leqslant \tau \leqslant n) = \Omega\};$$

$$\mathscr{M}_g^{m,n} = \{\tau \in \mathscr{M}^{m,n} : E_x g^-(\xi_\tau) < \infty \ (\forall\, x \in E)\};$$

$$S_{m,n}(x) = \sup\{E_x g(\xi_\tau) : \tau \in \mathscr{M}_g^{m,n}\};$$

$$\gamma_{m,n}(x,\omega) = \text{ess sup}\{E_x(g(\xi_\tau)/\mathscr{F}_m) : \tau \in \mathscr{M}_g^{m,n}\}.$$

定理 5.4.2 设 $g \in B_1$，则对$\forall\ 0 \leqslant m \leqslant n$，有

(i) $S_{m,n}(x) = E_x S_{n-m}(\xi_m)(\forall\ x \in E)$；

(ii) $\gamma_{m,n}(x,\omega) = S_{n-m}(\xi_m)(\text{mod}P_x)(\forall\ x \in E)$；

(iii) 令 $\tau_{m,n}^* = \min\{m \leqslant k \leqslant n : S_{n-k}(\xi_k) = g(\xi_k)\}$，$\text{mid}\{\varnothing\} = n$，必有

(a) $\tau_{m,n}^* \in \mathscr{M}_g^{m,n}$；

(b) $E_x(g(\xi_{\tau_{m,n}^*})/\mathscr{F}_m) = \gamma_{m,n}(x,\omega)(\text{mod}P_x)(\forall\ x \in E)$；

(c) $S_{m,n}(x) = E_x g(\xi_{\tau_{m,n}^*})(\forall\ x \in E)$.

证 设 θ 为平移算子，θ^m 为 m 步平移算子，则

$$f(\xi_k : m \leqslant k \leqslant n) = \theta^m f(\xi_0, \xi_1, \cdots, \xi_{n-m}).$$

然后，利用 ξ 的齐次 Markov 性即可将此定理转化成定理 5.4.1.

注 3 上述定理表达动态规划中的"向后递归原理". 它可以作为寻找最优停止规则的重要手段. 工作流程如下：

第一步，在 $\mathscr{M}_g^{n,n}$ 中解最优停止问题. 因为它只含一个元素 $\tau = n$. 因此，$\tau = n$ 即为最优停时. 于是

$$S_{n,n}(x) = E_x g(\xi_n).$$

第二步，在类 $\mathscr{M}_g^{n-1,n}$ 中解同一问题. 令

$$\tau_{n-1,n}^* = \begin{cases} n-1, & g(\xi_{n-1}) \geqslant E_x(g(\xi_n)/\mathscr{F}_{n-1}), \\ n, & g(\xi_{n-1}) < E_x(g(\xi_n)/\mathscr{F}_{n-1}). \end{cases}$$

则

$$\begin{aligned} E_x g(\xi_{\tau_{n-1,n}^*}) &= E_x(g(\xi_{n-1}) 1_{(\tau_{n-1,n}^* = n-1)}) + E_x(g(\xi_n) \cdot 1_{(\tau_{n-1,n}^* = n)}) \\ &= E_x Q g(\xi_{n-1}) = E_x S_{n-(n-1)}(\xi_{n-1}) \\ &= S_{n-1,n}(x) \quad (\text{按定理 } 5.4.2) \end{aligned}$$

由此即知：$\tau_{n-1,n}^*$ 为最优停时.

第三步，重复第二步中的处理方法.便可一步一步地作出最优停时列：$\tau_{n-2,n}^*,\cdots,\tau_{0,n}^*$，最后的 $\tau_{0,n}^*$ 即为 $(0,S_n)$-最优停时.

例5.4.1 挑选秘书问题.

这个问题的目标在于从 n 个对象中选出最好的秘书.作下列假设：

（Ⅰ）n 个对象按品质好坏编号为 $1,2,\cdots,n$."1"代表最好的.

（Ⅱ）n 个对象在时刻 $1,2,\cdots,n$ 内以一个随机序列独立到达.这 n 个对象的排列是等可能的.

（Ⅲ）任意两个对象必有一个比另一个好，即使不知道他们的编号.

（Ⅳ）在知道某个有序号的对象之后，要么拒绝（但不能重置），要么接受（挑选过程终止）.

在上述假设下，要考虑的问题是：以最大概率选出最好的对象.下面首先给出这个问题的数学模型，设

$$\Omega=\{\omega:\omega=(\omega_1,\cdots,\omega_n),\omega_k\in \mathrm{N}_n,k\in \mathrm{N}_n,\omega_k\neq\omega_l(k\neq l)\},$$
$$\mathrm{N}_n=\{1,2,\cdots,n\}.$$

显然，Ω 中有 $n!$ 个元素.按假设（Ⅱ），可在 Ω 上定义概率分布

$$P(\{\omega\})=\frac{1}{n!}\ (\omega\in\Omega);$$

$$Y_k(\omega)=\sum_{i=1}^{k}I_{i,k}(\omega),\text{其中 } I_{i,k}(\omega)=\begin{cases}1, & \omega_i\leqslant\omega_k,\\ 0, & \omega_i>\omega_k.\end{cases}$$

$$\mathscr{F}_k=\sigma(Y_1,Y_2\cdots,Y_k)\ (k\in \mathrm{N}_n).$$

变量 $Y_k(\omega)$ 表示在样本 ω 的分量中前 $k-1$ 个的品质有 $(Y_k(\omega)-1)$ 个比第 k 个好，即是说到时刻 k，有 k 个应征者，比第 k 个好的有 (Y_k-1) 个.

命题5.4.1 $P(\{Y_k(\omega)=i\})=\dfrac{1}{k}\ (\forall\ 1\leqslant i\leqslant k\leqslant n)$.

证 令 $A_i^k=\{Y_k(\omega)=i\}$，则 $A_i^k\cap A_j^k=\varnothing(i\neq j)$，且 $\Omega=\bigcup\limits_{i=1}^{k}A_i^k$.设 $\|A_i^k\|$ 表示 A_i^k 中元素个数，则此命题等价于

$$\| A_i^k \| = \| A_j^k \| \ (1 \leqslant i, j \leqslant k).$$

不失一般性，证明：$\| A_1^k \| = \| A_2^k \|$ $(2 \leqslant k)$.

设 $\omega \in A_1^k$，则 $\omega_i > \omega_k (\forall\ 1 \leqslant i \leqslant k-1)$. 设 $i_0(\omega) \in N_{k-1}$，且满足要求：$\omega_j \geqslant \omega_{i_0(\omega)} (\forall\ j \in N_{k-1})$. 定义映射：

$$\varphi(i,k,\omega;j) = \begin{cases} \omega_k, & j = i, \\ \omega_i, & j = k, \\ \omega_j, & j \neq i \text{ 或 } k \end{cases} \qquad (\omega \in \Omega, i、j、k \in N_n).$$

令 $\omega'_j = \varphi(i_0(\omega), k, \omega; j) (j \in N_n)$，则 $\omega' \in A_2^k$. 设 $\omega, \bar{\omega} \in A_1^k$，且 $\omega \neq \bar{\omega}$. 令

$$\omega'_j = \varphi(i_0(\omega), k, \omega; j), \bar{\omega}'_j = \varphi(i_0(\bar{\omega}), k, \bar{\omega}; j) \ (j \in N_n).$$

则 $\omega', \bar{\omega}' \in A_2^k$，且易证：$\omega' \neq \bar{\omega}'$.

上述分析表明：$\| A_1^k \| \leqslant \| A_2^k \|$，利用映射 φ 做同样的分析即可证明：$\| A_2^k \| \leqslant \| A_1^k \|$. 故 $\| A_1^k \| = \| A_2^k \|$.

命题 5.4.2 $Y = (Y_k, 1 \leqslant k \leqslant n)$ 为独立非齐次 Markov 链.

证 设 $i_l \in N_n$. 注意，$\{Y_1 = 1\} = \Omega$. 于是，

$$P(Y_l = i_l, 1 \leqslant l \leqslant k) = P(Y_l = i_l, 2 \leqslant l \leqslant k).$$

假如：$i_l \leqslant l$，则利用命题 5.4.1 之证明中构作映射的办法可以推出：

$$\begin{aligned} & P(Y_2 = 1, Y_l = i_l, 3 \leqslant l \leqslant k) \\ & = P(Y_2 = 2, Y_l = i_l, 3 \leqslant l \leqslant k) \\ & \Rightarrow P(Y_2 = i_2, Y_l = i_l, 3 \leqslant l \leqslant k) \\ & = \frac{1}{2} P(Y_l = i_l, 3 \leqslant l \leqslant k) \\ & = P(Y_2 = i_2) \cdot P(Y_l = i_l, 3 \leqslant l \leqslant n) \\ & \Rightarrow P(Y_l = i_l, 1 \leqslant l \leqslant k) = \prod_{l=1}^{k} P(Y_l = i_l). (5.4\text{-}1). \end{aligned}$$

如果有某 $l < i_l$. 则 $\{Y_l = i_l\}$ 为不可能事件. 从而，(5.4-1)式仍成立. 综合起来即得结论：Y 是独立的，独立序列必为 Markov 链，按命题 5.4.1，

$$P(Y_k = i_k) = \frac{1}{k} \neq \frac{1}{j} = P(Y_j = i_j)\ (k \neq j),$$

其中 $i_k \leqslant k, i_j \leqslant j$. 于是,$Y$ 是非齐次的.

命题 5.4.3 设 η_0 为取整数值的 r.v. 且独立于 Y;$\eta_k = \eta_0 + k$, $Y_k = 0\ (\forall\ k \leqslant 0)$. 令 $\xi_k = (\eta_k, Y_{\eta_k})$, $\xi = (\xi_k, 1 \leqslant k \leqslant n)$, $\widetilde{\mathscr{F}}_k = \sigma(\eta_0, Y_1, \cdots, Y_k)$,则 ξ 关于$(\widetilde{\mathscr{F}}.)$为齐次 Markov 链(二维的).

证 利用独立性,易证 ξ 的 Markov 性. 下证齐次性,设 $u = (u_1, u_2)$, $v = (v_1, v_2)$, $u, v \in \mathrm{N}_n \times \mathrm{N}_n$,则

$$\begin{aligned}
P(\xi_{k+1} = u/\xi_k = v) &= P(\eta_{k+1} \\
&= u_1, Y_{u_1} = u_2/\eta_k = v_1, Y_{v_1} = v_2) \\
&= P(\eta_{k+1} = u_1/\eta_k = v_1)P(Y_{u_1} = u_2/Y_{v_1} = v_2) \\
&= \begin{cases} P(Y_{v_1+1} = u_2), & \text{当 } u_1 = v_1 + 1 \text{ 时}, \\ 0, & \text{否则}. \end{cases}
\end{aligned}$$

此结果的右边与 k 无关. 这表明:ξ 具有齐次性.

命题 5.4.4 $P(\omega_k = 1/\mathscr{F}_k) = P(\omega_k = 1/Y_k) = \begin{cases} \dfrac{k}{n}, & Y_k = 1, \\ 0, & Y_k > 1. \end{cases}$

证 显然,$\{\omega_k = 1\} \subset \{Y_k = 1\}$. 因此,

$$\begin{aligned}
P(\omega_k = 1/Y_k = 1) &= \frac{P(\omega_k = 1, Y_k = 1)}{P(Y_k = 1)} \\
&= \frac{P(\omega_k = 1)}{P(Y_k = 1)} = \frac{k}{n}.
\end{aligned}$$

注意,$\{\omega_k = 1\} \cap \{Y_k > 1\} = \varnothing$, $P(Y_k > 1) = \dfrac{k-1}{k} > 0$. 因此,

$$P(\omega_k = 1/Y_k > 1) = \frac{P(\omega_k = 1, Y_k > 1)}{P(Y_k > 1)} = 0.$$

按假设(Ⅱ),$\{\omega_k = 1\}$独立于$\{Y_l = i_l\}(\forall\ l \in \mathrm{N}_{k-1})$. 故

$$P(\omega_k = 1/\mathscr{F}_k) = P(\omega_k = 1/Y_k).$$

命题 5.4.5 设 $g(\xi_k) = P_z(\omega_k = 1/\widetilde{\mathscr{F}}_k)(\forall\ k \in \mathrm{N}_n)$,其中 z 为

命题 5.4.3 中的二维齐次 Markov 链 ξ 的初始状态. 若 $z=(1,1)$，则

(i) $g(\xi_k)=P(\omega_k=1/\mathscr{F}_k)$ $(\forall\ k\in \mathrm{N}_n)$；

(ii) $P(\omega_\tau=1)=E_z g(\xi_\tau)$ $(\forall\ \tau\in\mathscr{T}_n)$，其中

$$\mathscr{T}_n=\{\tau:\tau\ \text{为取值于}\ \mathrm{N}_n\ \text{的}(\widetilde{\mathscr{F}}.)\ \text{时}\}.$$

证 由 $z=(1,1)$ 知：$\eta_1=1$. 从而，$\eta_0=0$，$(\mathscr{F}.)=(\mathscr{F}.)$，$P_z=P$，$\tau$ 亦为 $(\mathscr{F}.)$ 时，故

$$P(\omega_k=1/\mathscr{F}_k)=P(\omega_k=1/Y_k)=P_z(\omega_k=1/\xi_k)=g(\xi_k);$$

$$P(\omega_\tau=1)=P_z(\omega_\tau=1)=E_zE_z(1_{\{\omega_\tau=1\}}/\widetilde{\mathscr{F}}_\tau)=E_zg(\xi_\tau).$$

这里用到 $1_{\{\omega_k=1\}}$ 在 P_z 下独立于 $(\xi_l:1\leqslant l\leqslant k-1)$ 这一事实.

注 如果 $Z\neq(1,1)$，上述命题不必成立.

现在，我们可以叙述相应的最优停止问题. 设 $g(z)$ $(z\in \mathrm{N}_n\times \mathrm{N}_n)$ 为目标函数，ξ 作为观测过程，则最优停时 τ^* 满足如下关系：

$$E_{(1,1)}g(\xi_{\tau^*})=\sup\{E_{(1,1)}g(\xi_\tau):\tau\in\mathscr{T}_n\}.$$

按命题 5.4.5，此式的等价形式为

$$P(\omega_{\tau^*}=1)=\sup\{P(\omega_\tau=1):\tau\in\mathscr{T}_n\}.$$

这里 $P(\omega_{\tau^*}=1)$ 正是需要找的概率. 现将此表述标准化. 令 $\gamma_k=\xi_{k+1}$，则 $\gamma=(\gamma_k,0\leqslant k\leqslant n-1)$ 为齐次 Markov 链，起始时刻为 0. 目标函数为 $g(z)$. 按定理 5.4.1 和 5.4.2，有

$$S_{n-1-k}(z)=g(z)\vee TS_{n-2-k}(z)\quad(\forall\ z\in \mathrm{N}_n\times \mathrm{N}_n).$$

$$\begin{aligned}TS_{n-2-k}(\gamma_k)&=E_{(1,1)}(S_{n-2-k}(\gamma_{k+1})/\gamma_k)\\&=E_{(1,1)}S_{n-2-k}(\gamma_{k+1})\quad(\gamma\ \text{关于}\ P_{(1,1)}\ \text{独立})\\&=\sum_{i=1}^{k+2}S_{n-2-k}(k+2,i)P(Y_{k+2}=i)\\&=\frac{1}{k+2}\sum_{i=1}^{k+2}S_{n-2-k}(k+2,i)\end{aligned}$$

$$\Rightarrow S^{(k+1)}(y)=g(k+1,y)\vee\frac{1}{k+2}\sum_{i=1}^{k+2}S^{(k+2)}(i),$$

其中 $$S^{(k+1)}(y) \triangleq S_{n-1-k}(k+1, y).$$

令 $m=k+1$,则可得如下向后递归方程

$$\begin{cases} S^{(m)}(y) = g(m,y) \vee \dfrac{1}{m+1}\sum_{i=1}^{m+1} S^{(m+1)}(i) \quad ((m,y) \\ \qquad\qquad\qquad\qquad\qquad\qquad \in \mathrm{N}_{n-1} \times \mathrm{N}_n); \qquad (5.4\text{-}2) \\ S^{(n)}(y) = g(n,y) = \begin{cases} 1, & y=1, \\ 0, & y>1. \end{cases} \quad (y \in \mathrm{N}_n). \qquad (5.4\text{-}3) \end{cases}$$

命题 5.4.6 设 $m^* = m^*(n) \leqslant n$ 由下式确定:

$$\sum_{k=n-1}^{m^*} \frac{1}{k} \leqslant 1 < \sum_{k=n-1}^{m^*-1} \frac{1}{k}. \qquad (5.4\text{-}4)$$

则 $$S^{(k)}(1) = S^{(k)}(j)(\forall\, j \in \mathrm{N}_n, k \in \mathrm{N}_{m^*-1}),$$

$$S^{(m)}(y) = \begin{cases} \dfrac{m}{n}, & y=1, \\ \dfrac{m}{n}\sum_{k=n-1}^{m} \dfrac{1}{k}, & y>1 \end{cases} \quad (m^* \leqslant m \leqslant n-1).$$

证 按(5.4-2)和(5.4-3)式,有

$$S^{(n-1)}(1) = g(n-1,1) \vee \frac{1}{n}\sum_{i=1}^{n} S^{(n)}(i)$$

$$= g(n-1,1) \vee \frac{1}{n} = \frac{n-1}{n}, \qquad (5.4\text{-}5)$$

$$S^{(n-1)}(j) = g(n-1,j) \vee \frac{1}{n}\sum_{i=1}^{n} S^{(n)}(i)$$

$$= g(n-1,j) \vee \frac{1}{n}$$

$$= \frac{1}{n} = \frac{n-1}{n} \cdot \frac{1}{n-1} (j>1), \qquad (5.4\text{-}6)$$

按(5.4-2)和(5.4-5)式及(5.4-6)式,得

$$S^{(n-2)}(1) = g(n-2,1) \vee \frac{1}{n-1}\sum_{i=1}^{n-1} S^{(n-1)}(i)$$

$$= \frac{n-2}{n} \vee \left(\frac{1}{n-1}\left(\frac{n-1}{n} + \frac{n-2}{n}\right)\right) = \frac{n-2}{n},$$

$$S^{(n-2)}(j)=g(n-2,j)\vee\left(\frac{1}{n-1}\left(\frac{n-1}{n}+\frac{n-2}{n}\right)\right)$$

$$=\frac{1}{n-1}\left(\frac{n-1}{n}+\frac{n-2}{n}\right)=\frac{n-2}{n}\left(\frac{1}{n-1}+\frac{1}{n-2}\right)$$

$$(j>1).$$

按此方式后归至 m^*-1,则得

$$S^{(m^*-1)}(1)=g(m^*-1,1)\vee\frac{1}{m^*}\sum_{i=1}^{m^*}S^{(m^*)}(i)$$

$$=\frac{m^*-1}{n}\sum_{k=n-1}^{m^*-1}\frac{1}{k};\tag{5.4-7}$$

$$S^{(m^*-1)}(j)=g(m^*-1,j)\vee\frac{1}{m^*}\sum_{i=1}^{m^*}S^{(m^*)}(i)$$

$$=\frac{m^*-1}{n}\sum_{k=n-1}^{m^*-1}\frac{1}{k}(j>1).\tag{5.4-8}$$

按(5.4-4),(5.4-7)和(5.4-8)式,可得出

$$S^{(m^*-1)}(1)=S^{(m^*-1)}(j)\ (j\in \mathrm{N}_n);$$

$$S^{(k)}(1)=S^{(k)}(j)\ (j\in N_n,k\in \mathrm{N}_{m^*-1}).$$

命题 5.4.7 $P(\omega_{\tau^*}=1)=E_{(1,1)}g(\xi_{\tau^*})=\frac{m^*-1}{n}\sum\limits_{k=m^*-1}^{n-1}\frac{1}{k}$,其中 m^* 由(5.4-4)式决定

$$\tau^*=\min\{m^*\leqslant m\leqslant n:Y_m=1\}.\min(\varnothing)=n.$$

证 按定理 5.4.1,最优停时 τ^* 为

$$\tau^*=\min\{m\in N_n:S^{(m)}(Y_m)=g(\xi_m)\}.$$

按命题 5.4.6,有

$$S^{(m)}(1)\begin{cases}=\dfrac{m}{n}, & \text{当 } m\geqslant m^* \text{ 时},\\ >\dfrac{m}{n}, & \text{当 } m<m^* \text{ 时}.\end{cases}$$

$$S^{(m)}(y)>0\ (\forall\ (m,y)\in \mathrm{N}_{n-1}\times \mathrm{N}_n).$$

$$S^{(n)}(y) = g(n,y). \quad (y \in \mathrm{N}_n).$$

按 g 的定义,有

$$g(m,y) = \begin{cases} \dfrac{m}{n}, & y = 1, \\ 0, & y > 1. \end{cases} \quad (\forall\ (m,y) \in \mathrm{N}_n \times \mathrm{N}_n).$$

于是,τ^* 的表示可转化为

$$\tau^* = \min\{m^* \leqslant m \leqslant n : Y_m = 1\}, \min(\varnothing) = n$$

$$\Rightarrow E_{(1,1)} g(\xi_{\tau^*}) = ES^{(\tau^*)}(1) = \frac{m^* - 1}{n} \sum_{k=m^*-1}^{n-1} \frac{1}{k}.$$

命题 5.4.8 $P(\omega_{\tau^*} = 1) \xrightarrow[n\to\infty]{} \dfrac{1}{\mathrm{e}}.$

证 按(5.4-4)式及积分求和公式可知

$$m^*(n) \sim \frac{n}{\mathrm{e}} \text{ 当 } n \uparrow \infty \text{ 时}.$$

按命题 5.4.7,即得这里的结论.

§5.5 过剩函数与最小过剩主部

前一节讨论观测过程为有限齐次 Markov 链 ξ. 现在考虑可列的情形. 设目标函数为 g,则对 $\forall\, n \in \mathrm{N}$,最优报酬函数 $S_n(x) = Q^n g(x)$,且

$$S_n(x) = g(x) \vee TS_{n-1}(x) \text{ (见定理 5.4.1)}.$$

如果让 $n \to \infty$,最优停止理论如何展开? 本节就讨论这个问题. 在 §5.2 中,一般问题涉及到最大广义下鞅. 相反,本节将涉及最小广义上鞅.

定义 5.5.1 设 $f \in B$. 若 $Tf(x) = E_x f(\xi_1)$ 对 $\forall\, x \in E$ 有定义,且 $Tf(x) \leqslant f(x) (\forall\, x \in E)$,则称 f 为关于 ξ 的过剩函数.

定义 5.5.2 设 $f \in B$ 为过剩函数,$g \in B$. 若

$$g(x) \leqslant f(x) \ (\forall\ x \in E),$$

则称 f 为 g 的过剩主部. 若对 g 的任意过剩主部 u 还有 $f(x) \leqslant$

$u(x)(\forall\ x\in E)$，则称 f 为 g 的最小过剩主部.

如果 $\quad E_x(\sup\limits_n f^{\pm}(\xi_n))<\infty(\forall\ x\in E)$，

则说 f 满足"条件 $A^{\pm}$".

如果 $E_x f^-(\xi_\infty)<\infty(\forall\ x\in E)$，其中 $f(\xi_\infty)=\varlimsup\limits_{n\to\infty} f(\xi_n)$，

则说 f 满足"条件 a^-".

定义下列符号：

$$B^{\pm}=\{g\in B: g \text{ 满足条件 } A^{\pm}\},\ B^{-,+}=B^-\cup B^+,$$

$$B(a^-)=\{g\in B: g \text{ 满足条件 } a^-\}.$$

显然，常数为过剩函数. 下面讨论过剩函数的主要性质：

P. 1 若 f,g 为非负过剩函数，a、b 为非负实数，则 $af+bg$ 亦为过剩函数.

P. 2 若 $(f_n,n\geqslant 1)$ 为 B_1 中的非降过剩函数列，则极限 $f(x)=\lim\limits_{n\to\infty} f_n(x)$ 亦为过剩函数.

事实上，$-\infty<Tf_n(x)\leqslant f_n(x)\Rightarrow f_n\in B$. 注意 $f_1(x)\leqslant f_n(x)$. 按单调收敛定理，有

$$Tl_n \lim_{n\to\infty} f_n(x)) = \lim_{n\to\infty} Tf_n(x) \leqslant \lim_{n\to\infty} f_n(x)(\forall\ x\in E).$$

按定义 5. 5. 1，极限 f 为过剩函数.

P. 3 设 f 为过剩函数，且 $E_x f^-(\xi_n)<\infty(\forall\ n\in\mathrm{N})$，则 $(f(\xi_n),\mathscr{F}_n,n\in\mathrm{N})$ 为关于 P_x 的广义上鞅$(\forall\ x\in E)$.

事实上，按 ξ 的齐次 Markov 性，对$\forall\ x\in E$，有

$$\begin{aligned} E_x(f(\xi_{n+1})/\mathscr{F}_n) &= E_x(f(\xi_{n+1})/\xi_n) \\ &= E_{\xi_n} f(\xi_1) \leqslant f(\xi_n)\ (\mathrm{mod} P_x). \end{aligned}$$

P. 4 设 $f\in B^-$，且为过剩函数. 令

$$f_m(x) = T^m f(x) = E_x f(\xi_m),\ T^m = T\circ T\circ\cdots\circ T,$$

则 $(f_m,m\geqslant 1)$ 为过剩函数列，且对$\forall\ x\in E$，$(f_m(x),m\geqslant 1)$ 单调不增.

事实上，按 ξ 的齐次 Markov 性，有

$$E_x f(\xi_m) = E_x E_x(f(\xi_m)/\mathscr{F}_{m-1})$$

$$=E_x E_{\xi_{m-1}} f(\xi_1) = E_x T f(\xi_{m-1})$$

$$\Rightarrow E_x f(\xi_m) = T^m f(x).$$

按条件,$Tf(x) \leqslant f(x)$. 两边连续作用算子 T,得

$$T^{m+1} f(x) \leqslant T^m f(x) \Rightarrow f_m \downarrow .$$

P. 5 若 $f, g \in B_1$,且为过剩函数,则 $f \wedge g \in B_1$,且为过剩函数.

事实上,令 $h = f \wedge g$,则

$$h^-(x) = f^-(x) \vee g^-(x)$$

$$\Rightarrow E_x h^-(\xi_1) \leqslant E_x f^-(\xi_1) + E_x g^-(\xi_1) < \infty \Rightarrow h \in B_1.$$

显然,有下列关系成立:

$$E_x h(\xi_1) \leqslant E_x f(\xi_1) \wedge E_x g(\xi_1) = Tf(x) \wedge Tg(x)$$

$$\leqslant f(x) \wedge g(x)$$

$$\Rightarrow Th(x) \leqslant h(x) \ (\forall\ x \in E).$$

P. 6　如果过剩函数 f 满足条件:

$$\sup_n E_x f^-(\xi_n) < \infty \ (\forall\ x \in E),$$

则极限 $f(\xi_\infty) = \lim_{n \to \infty} f(\xi_n)$ $(\mathrm{mod} P_x)$存在,且$(\mathrm{mod} P_x)$有限或$+\infty$.

事实上,按 P. 3,$(f(\xi_n), n \in N)$关于 P_x 为广义上鞅. 按定理 4. 6. 2,结论真.

注　如果$\exists\ \eta \in L'(\Omega, \mathscr{F}, P_x)(\forall\ x \in E)$,使得

$$f(\xi_n) \geqslant E_x(\eta / \mathscr{F}_n) \ (\mathrm{mod} P_x) \ (\forall\ x \in E, n \in \mathrm{N}),$$

则必有$\sup_n E_x f^-(\xi_n) < \infty \ (\forall\ x \in E)$. 这里给出的条件表明:$(f(\xi_n), n \in \mathrm{N})$受一个一致可积鞅的下控.

P. 7 若 $f \in B^-$,且为过剩函数,则对$\forall$ 停时 σ 和 $\tau, \tau \geqslant \sigma (\mathrm{mod} P_x)(\forall\ x \in E)$,有

$$E_x(f(\xi_\tau) / \mathscr{F}_\sigma) \leqslant f(\xi_\sigma) \ (\mathrm{mod} P_x)(\forall\ x \in E).$$

事实上,$(-f(\xi_n), n \in N)$满足定理 4. 6. 5 的要求. 这是按 P. 3 和本性质的假设得到的.

P. 8　设过剩函数 $f \in B^-$,

$$\sigma = \inf\{n \in N: \xi_n \in B\}, \inf\{\varnothing\} = \infty \ (B \in \mathscr{E}).$$

则 $f_B(x) = E_x f(\xi_\sigma)$ 亦为过剩函数.

事实上,令 $\tau = \theta\sigma$,其中 θ 为一步平移算子. 则 τ 亦为停时,且 $\sigma \leqslant \tau$. 按 P. 7,可得

$$\begin{aligned} Tf_B(x) &= E_x E_{\xi_1} f(\xi_\sigma) = E_x E_x(f(\xi_{\sigma+1})/\mathscr{F}_1) \\ &= E_x f(\xi_{\theta\sigma}) = E_x E_x(f(\xi_\tau)/\mathscr{F}_\sigma) \\ &\leqslant E_x f(\xi_\sigma) = f_B(x). \end{aligned}$$

P. 9　设 $g \in B_1$,v 为 g 的最小过剩主部,则

$$v(x) = g(x) \vee Tv(x) \ (\forall \ x \in E).$$

事实上,按假设,有 $g(x) \leqslant v(x)$,$Tg(x) \leqslant Tv(x)$,令 $v_1(x) = g(x) \vee Tv(x)$. 则

$$\begin{aligned} & g(x) \leqslant v_1(x), Tv(x) \leqslant v_1(x), \\ & v_1(x) \leqslant v(x) \vee Tv(x) = v(x) \\ \Rightarrow & -\infty < Tg(x) \leqslant Tv_1(x) \leqslant Tv(x) \leqslant v_1(x). \end{aligned}$$

于是,$v_1 \in B_1$,$g(x) \leqslant v_1(x)$,且 v_1 为 g 的过剩函数. 按 v 的最小性,得 $v(x) \leqslant v_1(x)$ ($\forall \ x \in E$). 故 $v = v_1$.

P. 10　设 $g \in B_1$,算子 Q 定义如下:

$$Qg(x) = g(x) \vee Tg(x) \ (\forall \ x \in E).$$

则 g 的最小过剩主部 v 为

$$v(x) = \lim_{n\to\infty} Q^n g(x) \ (\forall \ x \in E).$$

事实上,按引理 5. 4. 1,$(Q^n g(x), n \in \mathrm{N})$ 单调不降. 按单调收敛定理,有 $Tv(x) \leqslant v(x)$. 因此,v 为 g 的过剩主部. 现设 f 为 g 的任意过剩主部,即 $g(x) \leqslant f(x)$,$Tf(x) \leqslant f(x)$. 则

$$\begin{aligned} & Qf(x) = f(x) \ (\forall \ x \in E) \\ \Rightarrow & f(x) = Q^n f(x) \geqslant Q^n g(x) \Rightarrow f(x) \geqslant v(x) \ (\forall \ x \in E). \end{aligned}$$

P. 11　设 $g \in B_1$,算子 $\tilde{Q}$ 为:$\tilde{Q}g(x) = \sup_k T^k g(x)$,$v$ 为 g 的最小过剩主部,则

$$v(x) = \lim_{n\to\infty} \tilde{Q}^n g(x) \ (\forall \ x \in E).$$

事实上，显然有 $Q^n g(x) \geqslant \max_{0\leqslant k\leqslant n} T^k g(x)$. 于是

$$v(x) = \lim_{n\to\infty} Q^n g(x) \geqslant \tilde{Q}\, g(x).\ (\forall\ x \in E)$$

$$\Rightarrow v(x) = Q^m v(x) \geqslant Q_0^m \tilde{Q}\, g(x)$$

$$\Rightarrow v(x) \geqslant \lim_{m\to\infty} Q_0^m \tilde{Q}\, g(x) \geqslant \tilde{Q}^2 g(x).$$

采用归纳法，便可得：

$$v(x) \geqslant \tilde{Q}^n g(x)\ (\forall\ x \in E, n \in \mathrm{N}).$$

按 $\tilde{Q}$ 之定义，有 $\tilde{Q}^n g(x) \uparrow$ 当 $n \uparrow \infty$ 时，且 $\tilde{Q}^n g \in B_1$.

令 $v_1(x) = \lim_{n\to\infty} \tilde{Q}^n g(x)$，则 $v \geqslant v_1$，且

$$Tv_1(x) = \lim_{n\to\infty} T\tilde{Q}^n g(x) \leqslant \lim_{n\to\infty} \tilde{Q}^{n+1} g(x) = v_1(x)\ (\forall\ x \in E).$$

即 v_1 为 g 的过剩主部，按 v 的最小性，得 $v = v_1$.

P. 12　设 $g \in B$，f 为 g 的过剩主部，且满足方程

$$f(x) = g(x) \vee Tf(x)\ (\forall\ x \in E).$$

令 $\tau_\varepsilon = \inf\{n \in N: f(\xi_n) \leqslant g(\xi_n) + \varepsilon\}, \inf\{\varnothing\} = \infty\ (\varepsilon \geqslant 0)$.

若对某固定的 $x \in E$，有 $f(x) < \infty$，则

$$E_x f(\xi_{n\wedge\tau_\varepsilon}) = f(x)\ (\forall\ n \in \mathrm{N}).$$

事实上，按条件，$|f(x)| < \infty$. 由 $E_x f(\xi_0) = f(x)$ 可知：$E_x(f(\xi_0) 1_{(\tau_\varepsilon > 0)})$ 和 $E_x(f(\xi_0) 1_{(\tau_\varepsilon = 0)})$ 均是有限的，且

$$f(x) = E_x(f(\xi_0) 1_{(\tau_\varepsilon > 0)}) + E_x(f(\xi_0) 1_{(\tau_\varepsilon = 0)}).$$

按 τ_ε 之定义，$f(\xi_0) > g(\xi_0)$ 于 $(\tau_\varepsilon > 0)$ 上. 于是，按假设，

$$f(\xi_0) = Tf(\xi_0) = E_x(f(\xi_1)/\mathscr{F}_0)\ \text{于}\ (\tau_\varepsilon > 0)\ \text{上}.$$

$$\Rightarrow f(x) = E_x(f(\xi_0) 1_{(\tau_\varepsilon = 0)}) E_x(f(\xi_1) 1_{(\tau_\varepsilon > 0)})$$

$$= E_x f(\xi_{1\wedge\tau_\varepsilon}) = E_x(f(\xi_{\tau_\varepsilon}) 1_{(\tau_\varepsilon \leqslant 1)})$$

$$= + E_x(f(\xi_1) 1_{(\tau_\varepsilon > 1)}) = E_x(f(\xi_{2\wedge\tau_\varepsilon})) = \cdots$$

$$= E_x f(\xi_{n\wedge\tau_\varepsilon}).\ (\forall\ n \in \mathrm{N}).$$

P. 13　设 $g \in B^+$，v 为 g 的最小过剩主部，则

$$\overline{\lim_{n\to\infty}} v(\xi_n) = \overline{\lim_{n\to\infty}} g(\xi_n)\ (\mathrm{mod} P_x)\ (\forall\ x \in E).$$

令 $\tau_\varepsilon = \inf\{n \in N: v(\xi_n) \leqslant g(\xi_n) + \varepsilon\}, \inf\{\varnothing\} = \infty$,则

$$P_x(\tau_\varepsilon < \infty) = 1 \ (\forall\ \varepsilon > 0, x \in E).$$

事实上,由 $g(x) \leqslant v(x)$ 可知

$$\overline{\lim} v(\xi_n) \geqslant \overline{\lim} g(\xi_n) \ (\mathrm{mod} P_x) \ (\forall\ x \in E).$$

下证反向关系. 为此,令

$$\psi_n = \sup\{g(\xi_k): k \geqslant n\}, \varphi(x) = E_x\psi_0, \varphi_n = E_x(\psi_n/\mathscr{F}_n).$$

则由 ξ 的齐次 Markov 性知

$$\varphi_n = \varphi(\xi_n), E_x\varphi(\xi_1) < \infty \ (\text{按 } g \in B^+),$$

$$g(x) \leqslant E_x\psi_0 = \varphi(x),$$

$$T\varphi(x) = E_x\varphi(\xi_1) = E_x\psi_1 \leqslant E_x\psi_0 = \varphi(x)$$

$$\Rightarrow v(x) \leqslant \varphi(x) \ (\forall\ x \in E)$$

$$\Rightarrow v(\xi_n) \leqslant \varphi(\xi_n) = \varphi_n.$$

对 $\forall\ m \leqslant n$,显然,有

$$E_x(\psi_m/\mathscr{F}_n) \geqslant E_x(\psi_n/\mathscr{F}_n) = \varphi_n \geqslant v(\xi_n). \ (\mathrm{mod} P_x)$$

$$\Rightarrow \overline{\lim_{n\to\infty}} E_x(\psi_m/\mathscr{F}_n) \geqslant \overline{\lim_{n\to\infty}} v(\xi_n) \ (\mathrm{mod} P_x) \ (\forall\ m \in \mathrm{N}, x \in E).$$

按 $g \in B^+$,$(E_x(\psi_m/\mathscr{F}_n), m \leqslant n)$ 关于 P_x 为广义鞅($\forall\ m \in \mathrm{N}, x \in E$),及定理 4.6.3(广义 Levy 定理),有
$\lim\limits_{n\to\infty} E_x(\psi_m/\mathscr{F}_n)$ $(\mathrm{mod} P_x)$ 存在,且

$$\lim_{n\to\infty} E_x(\psi_m/\mathscr{F}_n) \leqslant E_x(\psi_m/\mathscr{F}_\infty) = \psi_m (\mathrm{mod} P_x)$$

$$\Rightarrow \overline{\lim_{n\to\infty}} v(\xi_n) \leqslant \overline{\lim_{m\to\infty}} \psi_m = \overline{\lim_{n\to\infty}} g(\xi_n),$$

故命题的第一个结论成立.

命题中的第二结论等价于证明

$P_x(A) = 0$ ($\forall\ x \in E$),其中 $A = (\tau_\varepsilon = \infty)$.

按 $g \in B^+$,有

$$P_x(-\infty < \overline{\lim_{n\to\infty}} g(\xi_n) < \infty) = 1 \ (\forall\ x \in E).$$

按本命题的第一结论,有

$$P_x(\overline{\lim_{n\to\infty}} v(\xi_n) \geqslant \overline{\lim_{n\to\infty}} g(\xi_n) + \varepsilon) = 0. \text{ 即 } P_x(A) = 0.$$

P. 14　设 $g\in B$. 并定义算子 G 和 G^0：

$$G^0f(x)=f(x),Gf(x)=g(x)\vee Tf(x)\ (f\in B,x\in E).$$

若 $g\in B^+,\varphi(x)=E_x(\sup_n g(\xi_n))$，则对 $\forall\ x\in E$，有

(i) $G^{n+1}\varphi(x)\leqslant G^n\varphi(x)\ (\forall\ n\in \mathrm{N})$.

(ii) $\tilde{v}(x)=g(x)\vee T\tilde{v}(x)$，$\tilde{v}(x)=\lim_{n\to\infty}G^n\varphi(x)$.

(iii) $\overline{\lim_{n\to\infty}}\tilde{v}(\xi_n)=\overline{\lim_{n\to\infty}}g(\xi_n)\ (\mathrm{mod}P_x)$.

事实上，按 G,φ 之定义，有

$$\begin{aligned}G\varphi(x)&=g(x)\vee T\varphi(x)=g(x)\vee E_xE_{\xi_1}(\sup_n g(\xi_n))\\&=E_x(g(\xi_0))\vee E_x(\sup_{n\geqslant 1}g(\xi_n))\\&\leqslant E_x(g(\xi_0)\vee\sup_{n\geqslant 1}g(\xi_n))=\varphi(x)\end{aligned}$$

$$\Rightarrow G^n\varphi(x)\leqslant\varphi(x),G^{n+1}\varphi(x)\leqslant G^n\varphi(x).\ (\forall\ n\in\mathrm{N}),$$

即 $(G^n\varphi(x),n\in\mathrm{N})$ 单调不增，此即结论(i).

“$g\in B^+$”表明：$T\varphi(x)<\infty(\forall\ x\in E)$. 按 G 之定义，有

$$G^{n+1}\varphi(x)=g(x)\vee TG^n\varphi(x).$$

按单调收敛定理，可推出

$$\begin{aligned}\tilde{v}(x)&=\lim_{n\to\infty}G^n\varphi(x)=g(x)\vee\lim_{n\to\infty}TG^{n-1}\varphi(x)\\&=g(x)\vee T\tilde{v}(x)\ (\forall\ x\in E).\Rightarrow\text{(ii)}.\end{aligned}$$

由(ii)知，$\overline{\lim}\tilde{v}(\xi_n)\geqslant\overline{\lim}g(\xi_n)$. 另一方面，又有

$$\tilde{v}(\xi_n)\leqslant G^n\varphi(\xi_n)\leqslant\varphi(\xi_n)\leqslant E_x(\sup_{j\geqslant m}g(\xi_j)/\mathscr{F}_n)\ (\forall\ m\leqslant n)$$

$$\Rightarrow\overline{\lim_{n\to\infty}}\tilde{v}(\xi_n)\leqslant\overline{\lim_{m\to\infty}}g(\xi_m)\ (\mathrm{mod}P_x)\Rightarrow\text{(iii)}.$$

P. 15 设 $\tilde{\tau}_\varepsilon=\inf\{n\in\mathrm{N}:\tilde{v}(\xi_n)\leqslant g(\xi_n)+\varepsilon\},\inf\{\varnothing\}=\infty(\varepsilon>0)$，其中 $\tilde{v}$ 按上面的 P. 14 定义，则

(i) $\tilde{v}(x)\leqslant E_x\tilde{v}(\xi_{\tilde{\tau}_\varepsilon}),P_x(\tilde{\tau}_\varepsilon<\infty)=1(\forall\ x\in E)(g\in B^+)$;

(ii) $\tilde{v}(x)=E_x\tilde{v}(\xi_{\tilde{\tau}_\varepsilon}),\tilde{v}(x)=v(x)(\forall\ x\in E)(g\in B^{-,+})$，

其中 v 为 g 的最小过剩主部.

事实上，按 P. 14，$\tilde{v}$为 g 的过剩主部. 按 P. 12，有

$$\tilde{v}(x)=E_x\tilde{v}(\xi_{n\wedge\tilde{\tau}_\varepsilon})$$

$$=E_x(\tilde{v}(\xi_{\tilde{\tau}_\varepsilon})1_{(\tilde{\tau}_\varepsilon\leqslant n)})+E_x(\tilde{v}(\xi_n)1_{(\tilde{\tau}_\varepsilon>n)})(\forall\ n\in\mathrm{N}).$$

类似于 P. 13 中的处理方式便可得 $P_x(\tilde{\tau}_\varepsilon<\infty)=1$. 显然，按 P. 14. 有

$$E_x(\tilde{v}(\xi_n)1_{(\tilde{\tau}_\varepsilon>n)})\leqslant E_x(\sup_{j\geqslant n}g^+(\xi_j)1_{(\tilde{\tau}_\varepsilon>n)}).$$

于是，按 $g\in B^+$，Fatou 引理及单调收敛定理，可得：

$$\tilde{v}(x)\leqslant\overline{\lim_{n\to\infty}}E_x(\tilde{v}(\xi_{\tilde{\tau}_\varepsilon})1_{(\tilde{\tau}_\varepsilon\leqslant n)})$$

$$+\underline{\lim_{n\to\infty}}E_x(\sup_{j\geqslant n}g^+(\xi_j)1_{(\tilde{\tau}_\varepsilon>n)})$$

$$\leqslant E_x\tilde{v}(\xi_{\tilde{\tau}_\varepsilon}).\Rightarrow\text{(i)}.$$

$g\in B^{-,+}$表明：$g\in B^-$. 类似地证明可得

$$\tilde{v}(x)\geqslant E_x\tilde{v}(\xi_{\tilde{\tau}_\varepsilon}).$$

由于 $g\in B^+$，，此反向关系亦成立. 故

$$\tilde{v}(x)=E_x\tilde{v}(\xi_{\tilde{\tau}_\varepsilon})\ (\forall\ x\in E).$$

按 $\tilde{\tau}_\varepsilon$ 的有限性，v 的过剩性及 P. 7，得

$$\tilde{v}(x)=E_x\tilde{v}(\xi_{\tilde{\tau}_\varepsilon})\leqslant E_xg(\xi_{\tilde{\tau}_\varepsilon})+\varepsilon$$

$$\leqslant E_xv(\xi_{\tilde{\tau}_\varepsilon})+\varepsilon\leqslant v(x)+\varepsilon.\ (\forall\ \varepsilon>0).$$

$$\Rightarrow\tilde{v}(x)\leqslant v(x).$$

按 P. 14，$\tilde{v}$为 g 的过剩主部而 v 为 g 的最小过剩主部. 故 $\tilde{v}=v.\Rightarrow$ (ii).

注 P. 15(i)表明：当 $g\in B^+$时，$\tilde{v}$不一定为 g 的最小过剩主部. 但 v 和 $\tilde{v}$均满足如下方程：

$$\begin{cases}f(x)=g(x)\vee Tf(x)\ (\forall\ x\in E);\\ \overline{\lim_{n\to\infty}}f(\xi_n)=\overline{\lim_{n\to\infty}}g(\xi_n)\ (\mathrm{mod}P_x)\ (\forall\ x\in E),\end{cases}\tag{5.5-1}$$

即是说：上述方程当 $g\in B^+$时不必有唯一解（下一节有一个例将证实这一点）. 但 P. 15(ii)表明当 $g\in B^{-,+}$时，方程(5.5-1)有唯一

解.

定理 5.5.1 设 $g\in B^-$,则

(i) 报酬函数 $S(x)$ 是 g 的最小过剩主部;

(ii) 两种报酬函数 S 和 $\overline{S}$ 一致,即 $S=\overline{S}$;

(iii) S 满足方程(5.5-1);

(iv) $S(x)=\lim\limits_{n\to\infty}Q^n g(x)=\lim\limits_{b\to\infty}\lim\limits_{n\to\infty}Q^n g^{(b)}(x)$. $(\forall\ x\in E)$,

其中 $g^{(b)}(x)=g(x)\wedge b$;

(v) 当 $g\in B^{-,+}$ 时,有 $S(x)=E_xS(\xi_{\tau_\varepsilon})\leqslant E_xg(\xi_{\tau_\varepsilon})+\varepsilon$. $(\forall\ x\in E)$. $(\forall\ \varepsilon>0)$,

其中 $\tau_\varepsilon=\inf\{n\in \mathrm{N}:S(\xi_n)\leqslant g(\xi_n)+\varepsilon\},\inf\{\varnothing\}=\infty$.

$$P_x(\tau_\varepsilon<\infty)=1\ (\forall\ \varepsilon>0,x\in E).$$

证 首先考虑 $g\in B^{-,+}$ 的情形. 按 P.10,有

$$v(x)=\lim_{n\to\infty}Q^n g(x)\ (\forall\ x\in E).$$

且 v 为 g 的最小过剩主部. 按 P.7,有

$$v(x)\geqslant E_xv(\xi_\tau)\geqslant E_xg(\xi_\tau)\ (\forall\ \tau\in\overline{\mathscr{M}}_g,x\in E)$$

$$\Rightarrow v(x)\geqslant\sup\{E_xg(\xi_\tau):\tau\in\overline{\mathscr{M}}_g\}=\overline{S}(x)\geqslant S(x).$$

按 P.15(ii),有

$$v(x)=E_xv(\xi_{\tilde\tau_\varepsilon})\leqslant E_xg(\xi_{\tilde\tau_\varepsilon})+\varepsilon\leqslant S(x)+\varepsilon,$$

其中 $\tilde\tau_\varepsilon$ 按 P.15 中定义,综合起来,得

$$S(x)\leqslant\overline{S}(x)\leqslant v(x)\leqslant S(x)+\varepsilon\ (\forall\ \varepsilon>0,x\in E)$$

$$\Rightarrow S(x)=\overline{S}(x)=v(x)\ (\forall\ x\in E).$$

这表明:S 和 $\overline{S}$ 具有 v 的所有特性. 因此,$\tau_\varepsilon=\tilde\tau_\varepsilon$. 故当 $g\in B^{-,+}$ 时,定理中的所有结论成立.

现在设 $g\in B^-$,则对 $\forall\ b\geqslant0,g^{(b)}\in B^{-,+}$. 让 $v^{(b)}$ 为 $g^{(b)}$ 的最小过剩主部,则 $v^{(b)},b\geqslant0$,单调非降. 令

$$v^*(x)=\lim_{b\to\infty}v^{(b)}(x),$$

则由 $g\in B^-$ 及单调收敛定理可知

$$Tv^*(x)=\lim_{b\to\infty}Tv^{(b)}(x)\leqslant\lim_{b\to\infty}v^{(b)}(x)=v^*(x).$$

显然，$g\leqslant v^*$. 故 v^* 为 g 的过剩主部. 设 f 为 g 的任意过剩主部. 自然也为 $g^{(b)}$的过剩主部. 因此

$$f(x)\geqslant Qg^{(b)}(x),f(x)=Q^nf(x)\geqslant Q^{n+1}g^{(b)}(x)$$
$$\Rightarrow f(x)\geqslant v^{(b)}(x)\Rightarrow f(x)\geqslant v^*(x).$$

故 v^* 为 g 的最小过剩主部，即 $v^*=v$. 从而，有

$$v(x)=\lim_{b\to\infty}\lim_{n\to\infty}Q^ng^{(b)}(x),v(x)=\lim_{b\to\infty}v^{(b)}(x).$$

按前一步之证，有

$$S^{(b)}(x)=v^{(b)}(x)\Rightarrow S^*(x)=v^*(x)=v(x).$$

按单调收敛定理，有

$$E_xg(\xi_\tau)=\lim_{b\to\infty}E_xg^{(b)}(\xi_\tau)\ (\forall\ \tau\in\overline{\mathscr{M}}_p)$$
$$\Rightarrow\overline{S}(x)=\lim_{b\to\infty}\overline{S}^{(b)}(x)=\lim_{b\to\infty}S^{(b)}(x)=S^*(x),$$
$$S(x)=\lim_{b\to\infty}S^{(b)}(x)=S^*(x)$$

$\Rightarrow\overline{S}=S$. 按 P. 15 后之注，$S$ 满足方程(5. 5-1).

定理 5. 5. 2 设 $g\in B^{-,+}$，v 为 g 的最小过剩主部，

$\tau_\varepsilon=\inf\{n\in N:v(\xi_n)\leqslant g(\xi_n)+\varepsilon\},\inf\{\varnothing\}=\infty.\ (\varepsilon\geqslant0).$ 则

(i) τ_ε 为 g 的(ε,S)-最优停时$(\varepsilon>0)$；

(ii) τ_0 为 g 的$(0,\overline{S})$-最优停时；

(iii) 当 $\tau_0\in\mathscr{M}$ 时，τ_0 为$(0,S)$-最优停时；

(iv) 当状态空间 E 有限时，τ_0 为$(0,S)$-最优停时.

证　按 P. 15(i)，$P_x(\tau_\varepsilon<\infty)=1(\varepsilon>0,x\in E)$. 按定理 5. 5. 1，$\tau_\varepsilon$ 为 g 的(ε,S)-最优停时$(\varepsilon>0)$. 这就得(i).

按 P. 7，有(v 为 g 的最小过剩主部)

$$v(x)\leqslant E_xv(\xi_{\tau_0})(\because g\in B^+).$$

按 P. 13 及 τ_0 之定义，有

$$E_xv(\xi_{\tau_0})=E_x(v(\xi_{\tau_0})1_{(\tau_0<\infty)})+E_x(v(\xi_{\tau_0})1_{(\tau_0=\infty)})$$
$$=E_x(g(\xi_{\tau_0})1_{(\tau_0<\infty)})+E_x(g(\xi_\infty)1_{(\tau_0=\infty)})(\because g\in B^+)$$

$$=E_x g(\xi_{\tau_0}).$$

按定理 5.5.1(i)～(ii)，得

$$S(x)=\overline{S}(x)=v(x)\leqslant E_x g(\xi_{\tau_0}).$$

于是，定理 5.5.1 中的(ii)和(iii)成立.

最后考虑 E 为有限状态空间的情形. 设

$$\Gamma_\varepsilon=\{x\in E: v(x)\leqslant g(x)+\varepsilon\},$$

则$(\Gamma_\varepsilon,\varepsilon\geqslant 0)$具有下列性质：$\Gamma_0\subset\Gamma_\varepsilon$，$\Gamma_\varepsilon\downarrow\Gamma_0$；$\exists\ \varepsilon'>0$，使得当 $0\leqslant\varepsilon\leqslant\varepsilon'$ 时，$\Gamma_\varepsilon=\Gamma_0$.（事实上，若不然，即 $\varepsilon'>0$ 不存在. 于是，可选 $\varepsilon_1>0$ 及 $x_{\varepsilon_1}\in\Gamma_{\varepsilon_1}$，使得，$x_{\varepsilon_1}\notin\Gamma_0$. 再选 $0<\varepsilon_2<\varepsilon_1$ 及 $x_{\varepsilon_2}\in\Gamma_{\varepsilon_2}\setminus\{x_{\varepsilon_1}\}$，使得 $x_{\varepsilon_2}\notin\Gamma_0$. 这样就可构作一个点列$(x_{\varepsilon_k},k\geqslant 1)\not\subset\Gamma_0$，且 $x_{\varepsilon_k}\neq x_{\varepsilon_i}$ $(i\neq k)$. 这表明：E 不是有限的）. 于是，$\tau_\varepsilon=\tau_0(0\leqslant\varepsilon\leqslant\varepsilon')$. 故按定理 5.5.1(v)，有

$$P_x(\tau_0=\tau_\varepsilon<\infty)=1\ (\forall\ x\in E, 0<\varepsilon\leqslant\varepsilon').$$

按本定理的结论(iii)，得

$$S(x)=E_x g(\xi_{\tau_0})\ (\forall\ x\in E).\Rightarrow\text{(iv)}.$$

§ 5.6 例

例 5.6.1 随机游动的最优停止问题.

设$(\eta_n)_{n\geqslant 1}$为独立同分布的 r. v. 列，且 $P(\eta_n=1)=p$，$P(\eta_n=-1)=q$，这里 $p+q=1$. 状态空间 $\tilde{E}=\{0,\pm 1,\pm 2,\cdots\}$. ξ_0 为取值于 $\tilde{E}$，且独立于 η 的 r. v. $\xi_n=\sum\limits_{k=1}^{n}\eta_k+\xi_0$. $(n\geqslant 1)$. $\mathscr{F}_n=\sigma(\xi_0,\xi_1,\cdots,\xi_n)$. $(n\in\mathrm{N})$. 则 $\xi=(\xi_n)_{n\in\mathrm{N}}$为$(\Omega,\mathscr{F},P)$上的适应于$(\mathscr{F}_.)$的齐次 Markov 链，这里 $\mathscr{F}=\sigma(\bigcup\limits_n\mathscr{F}_n)$，并 ξ 以$(\tilde{E},\tilde{\mathscr{E}})$为状态空间，其中 $\tilde{\mathscr{E}}$表示 $\tilde{E}$中一切子集之全体. 概率 P_x 由 ξ 的转移概率和集中在点 $x\in\tilde{E}$的初始分布所决定.

目标函数 g：$g(x)=x^+(\forall\ x\in\tilde{E})$. 显然，$g\in B^-$.

命题5.6.1 若 $p\geqslant q$，则 $\overline{S}(x)\equiv\infty$，且 $(0,\overline{S})$-最优停时 $\tau_0\equiv\infty$.

证 $Tg(x)=E_x(x+\eta_1)^+=(x+1)^+p+(x-1)^+q\geqslant x^+$ ($\forall\ x\in\widetilde{E}$).

$$Qg(x)=g(x)\vee Tg(x)=Tg(x);$$
$$Q^n g(x)=T^n g(x)=E_x\xi_n^+;$$
$$\overline{S}(x)=\lim_{n\to\infty}Q^n g(x)=\lim_{n\to\infty}E_x\xi_n^+.$$

若 $p>q$，则

$$E_x\xi_n^+\geqslant E_x\xi_n=x+n\cdot P(p-q)\xrightarrow[n\to\infty]{}\infty.$$

若 $p=q$，则

$$E_x\xi_0^+=\sum_{k=1}^{n}k\cdot{}_x(\xi_n=k).$$

$$P_x(\xi_n=k)$$
$$=\begin{cases}\dfrac{1}{2^n}\dfrac{n!}{\left(\dfrac{n+k-x}{2}\right)!\left(\dfrac{n-k+x}{2}\right)!}, & \text{当 } n+k-x \text{ 为偶数时,}\\ 0, & \text{当 } n+k-x \text{ 为奇数时.}\end{cases}$$
$$(-n\leqslant k-x\leqslant n).$$

(参看第七章(Markov 链)例5.12). 假定 $x=0$. 利用 Sterling 公式：$n!\sim\sqrt{2\pi n}\left(\dfrac{n}{e}\right)^n\cdot\exp\left(\dfrac{\theta_n}{12n}\right)$ $(0<\theta_n<1)$ 当 $n\to\infty$ 时. 则可得如下渐近公式：

$$E_0\xi_n^+\sim C\cdot\frac{1}{\sqrt{n}}\sum_{k=1}^{\left[\frac{n}{2}\right]}k\xrightarrow[n\to\infty]{}\infty.$$

综上分析，得 $\overline{S}(x)\equiv\infty$. 显然，$\Gamma_0=\{x\in\widetilde{E}:\overline{S}(x)\leqslant g(x)\}=\varnothing$. 故 $\tau_0\equiv\infty$.

注1 当 $p>q$ 时，$\lim\limits_{n\to\infty}E_x\xi_n^+=\infty$. 因此，$E_x(\sup\limits_n\xi_n^+)=\infty$ ($\forall\ x\in\widetilde{E}$). 这表明：$g\notin B^+$. 但 $g\in B^-$. 可见，g 满足定理5.5.1的要求，但不满足定理5.5.2的要求.

命题5.6.2 $\overline{\lim_{n\to\infty}}\xi_n^+=\begin{cases}+\infty, & p>q,\\ 0, & p<q.\end{cases}$ $(\mathrm{mod}P_x)(\forall\ x\in\tilde{E})$.

证 显然，当 $p>q$ 时，ξ 为下鞅，从而，ξ^+ 为下鞅. 因此，按定理 4.6.2.，$\lim\limits_{n\to\infty}\xi_n^+$ 存在于集 A 上，有限或 $-\infty$. 取"$-\infty$"不可能.

于是

$$
\begin{aligned}
\{\xi_n^+\to\}&=\{\lim_{n\to\infty}\xi_n^+<\infty\}\\
&=A\triangleq\{\omega:\inf_k\sup_m E_x(\xi_m^+/\mathscr{F}_k)<\infty\}\\
&\qquad(\mathrm{mod}P_x)(\forall\ x\in\tilde{E}).
\end{aligned}
$$

注意：$E_x(\xi_m^+/\mathscr{F}_k)\geqslant\xi_k+(m-k)(p-q)(\forall\ m\geqslant k)$

$$
\begin{aligned}
&\Rightarrow\inf_k\sup_m E_x(\xi_m^+/\mathscr{F}_k)=\infty\ (\mathrm{mod}P_x)\\
&\Rightarrow A=\varnothing\ (\mathrm{mod}P_x).
\end{aligned}
$$

显然，$E_x(\sup\limits_n|\Delta\xi_n^+|)\leqslant 1$. 从而 $\xi^+\in\mathscr{U}^+$（见§4.8）. 按定理 4.8.1，

$$\{\overline{\lim_{n\to\infty}}\xi_n^+<\infty\}=\{\sup_n\xi_n^+<\infty\}=\{\xi_n^+\to\}=A=\varnothing(\mathrm{mod}P_x),$$

故 $\qquad\overline{\lim}\xi_n^+=\infty\ (\mathrm{mod}P_x)$.

现假定，$p<q$，则由 $C_n^i\leqslant C_n^{[\frac{n}{2}]}(0\leqslant i\leqslant n)$ 及 Sterling 公式可推出：

$$
\begin{aligned}
E_0\xi_n^+&=\sum_{k=1}^n kP_0(\xi_n=k)\leqslant\sum_{k=1}^n C_n^{[\frac{n}{2}]}k\cdot\left(\frac{p}{q}\right)^{\frac{k}{2}}(pq)^{\frac{n}{2}}\\
&\leqslant\alpha\frac{1}{\sqrt{n}}(4pq)^{\frac{n}{2}}\sum_{k=1}^n k\left(\frac{p}{q}\right)^{\frac{k}{2}}\xrightarrow[n\to\infty]{}0
\end{aligned}
$$

$$\Rightarrow\overline{\lim}\xi_n^+=\lim\xi_n^+=0\ (\mathrm{mod}P_0)\ \text{（按定理 4.5.2）}.$$

对于 $x\neq0$ 的情形. 可利用 ξ-x 代替 ξ 即可将问题转化成 $x=0$ 的情形.

命题5.6.3 设 $p<q,\Gamma_r=[r,\infty)\in\mathscr{E},\tau_r=\inf\{n\in\mathrm{N}:\xi_n\geqslant r\}$，$\inf(\varnothing)=\infty$.

则

$$P_x(\xi_{\tau_r} \geqslant r) = P_x(\tau_r < \infty)$$

$$= \begin{cases} \left(\dfrac{p}{q}\right)^{r-x}, & r > x, \\ 1, & r \leqslant x. \end{cases} \quad (\forall\, r、x \in E).$$

证　若 $r \leqslant x$，则 $x \in \Gamma_r$. 从而，$\tau_r = 0(\mathrm{mod} P_x)$. 于是

$$P_x(\xi_{\tau_r} \geqslant r) = P_x(\xi_0 \geqslant x) = 1.$$

现在，考虑 $r > x$. 若对 ξ 作平移变换 $\xi\text{-}x$，则问题变可转化为 $x=0$. 以下设 $x=0$. 令 $\beta_0 = P_0(\tau_r < \infty) = P_0(\xi_{\tau_r} \geqslant r) = P_0(\xi_{\tau_r} = r)$. 按命题 5. 6. 2，$\overline{\lim_{n\to\infty}}(-\xi_n) = +\infty$，$\overline{\lim_{n\to\infty}}\xi_n^+ = 0(\mathrm{mod} P_0)$. 显然，$\left(\left(\dfrac{p}{q}\right)^{\xi_n}\right)_{n\in N}$ 为鞅. 因此，按鞅的有界停时代替的不变性，有

$$1 = \left(\frac{p}{q}\right)^0 = E_0\left(\frac{p}{q}\right)^{-\xi_{n\wedge\tau_r}} (\forall\, n \in \mathrm{N}).$$

按 τ_r 之定义，$E_0\left[\left(\dfrac{p}{q}\right)^{-\xi_{\tau_r}} \cdot 1_{(\tau_r\leqslant n)}\right] = \left(\dfrac{p}{q}\right)^{-r} \cdot P_0(\tau_r \leqslant n) \xrightarrow[n\to\infty]{} \left(\dfrac{p}{q}\right)^{-r} \cdot \beta_0$. 令 $\theta_n = \left(\dfrac{p}{q}\right)^{-\xi_n}$，则由 $1 = \left(\dfrac{p}{q}\right)^{-r} \cdot P_0(\tau_r \leqslant n) + E_0\theta_n \cdot 1_{(\tau_r > n)}$ $(\forall\, n \in N)$ 知：$\lim\limits_{n\to\infty} E_0\theta_n \cdot 1_{(\tau_r>n)}$ 存在. 下面将证明：此极限为零.

令 $C^* = \lim\limits_{n\to\infty} E_0(\theta_n 1_{(\tau_r>n)})$. 并注意：$(\eta_n)_{n\in N}$ 的独立，及

$$(\tau_r > n) = \{\xi_l < r : 0 \leqslant l \leqslant n\}.$$

得

$$\begin{aligned} E_0(\theta_n \cdot 1_{(\tau_r>n)}) &= pE_0(\theta_{n-1} \cdot 1_{(\tau_r>n-1)}) \\ &\quad + qE_0(\theta_{n-1} \cdot 1_{(\tau_r>n-1)\cdot\{\xi_{n-1}<r-1\}}) \end{aligned}$$

$$\Rightarrow \lim_{n\to\infty} E_0(\theta_{n-1} \cdot 1_{(\tau_r>n-1)\cdot\{\xi_{n-1}<r-1\}}) = C^*.$$

反复这一推理过程，便可得

$$\lim_{n\to\infty} E_0(\theta_{n-m} 1_{(\tau_r>n-m)\cdot\{\xi_{n-m}<r-m\}}) = C^* (\forall\, m \in \mathrm{N})$$

$$\Rightarrow C^* \leqslant \left(\frac{p}{q}\right)^{m-r} (\forall\, m \in \mathrm{N}).$$

让 $m\to\infty$，便得 $C^*=0$. 由此结果便可得

$$1=\left(\frac{p}{q}\right)^{-r}\cdot\beta_0,\text{即 }\beta_0=\left(\frac{p}{q}\right)^{r}.$$

命题 5.6.4 设 $p<q$，$f_r(x)=E_xg(\xi_{\tau_r})=E_x\xi_{\tau_r}^+$，$r^*$ 为 $r\cdot\left(\frac{p}{q}\right)^{r}$ 在 $\tilde{E}$ 中的最大值点，则 $f^*(x)=E_xg(\xi_{\tau_{r^*}})$ 为 g 的一个过剩主部.

证 按命题 5.6.3，有

$$f_r(x)=\begin{cases}r^+\cdot P_x(\tau_r<\infty), & r\geqslant x,\\ x^+, & r<x\end{cases}$$

$$=\begin{cases}r^+\cdot\left(\frac{p}{q}\right)^{r-x}, & r\geqslant x,\\ x^+, & r<x.\end{cases}\quad(x\in\tilde{E}).$$

当 $x<0$ 时，由 $r^*\geqslant 0$ 知：$r^*\geqslant x$. 于是，

$$f^*(x)\geqslant 0=x^+=g(x).$$

当 $0\leqslant x\leqslant r^*$ 时，由 $x=r^*$ 为函数 $u(x)=r^*\cdot\left(\frac{p}{q}\right)^{r^*-x}-x$ 的最小值点知：$f^*(x)\geqslant x=x^+=g(x)$. 当 $r^*<x$ 时，

$$f^*(x)=x=x^+=g(x).$$

综上分析，得 $f^*(x)\geqslant g(x)$ $(\forall\, x\in\tilde{E})$.

为证 f^* 为 g 的一个过剩主部，剩下的问题就是证明：$f^*(x)\geqslant Tf^*(x)$ $(\forall\, x\in\tilde{E})$. 事实上

$$Tf^*(x)=p\cdot f^*(x+1)+q\cdot f^*(x-1).$$

记 $u(x)=f^*(x)-Tf^*(x)$，则

$$u(x)=(f^*(x)-f^*(x+1))p+(f^*(x)-f^*(x-1))\cdot q.$$

当 $r^*<x$ 时，

$$u(x)=p(x-(x+1))+q(x-(x-1))=-p+q>0.$$

当 $r^*>x$ 时，

$$u(x)=p\cdot r^*\cdot\left(\frac{p}{q}\right)^{r^*-x}\left(1-\frac{q}{p}\right)+q\cdot r^*\left(\frac{p}{q}\right)^{r^*-x}\left(1-\frac{p}{q}\right)=0.$$

为证 $u(x)\geqslant 0$，最后剩下证明：当 $x=r^*$ 时，$u(x)\geqslant 0$. 这时，

$$u(x)=p(r^*-(r^*+1))+q(r^*-r^*\cdot\frac{p}{q})$$

$$=-p+r^*\cdot(q-p)=r^*\left(q-\frac{1+r^*}{r^*}p\right).$$

于是,“$u(x)\geqslant 0$”$\Leftrightarrow$“$\frac{q}{p}\geqslant\frac{1+r^*}{r^*}$”. 这就导致讨论 $r\cdot\left(\frac{p}{q}\right)^r$ 的特性.

当 $r\in R_+$时,$r\cdot\left(\frac{p}{q}\right)^r$ 的极大值点为 $\tilde{r}=\left(\ln\left(\frac{q}{p}\right)\right)^{-1}$. 且有估计:$\frac{1+\tilde{r}}{\tilde{r}}=1+\ln\left(\frac{q}{p}\right)\leqslant\frac{q}{p}$. 若 $\tilde{r}\in E$,则取 $r^*=\tilde{r}$. 若 $\tilde{r}\notin E$,则取 $r^*=[\tilde{r}+1]$([·]表示取整),且有

$$\frac{1+r^*}{r^*}\leqslant\frac{1+\tilde{r}}{\tilde{r}}\leqslant\frac{q}{p}.$$

显然,r^* 是 $\tilde{E}$ 中使 $r\left(\frac{p}{q}\right)^r$ 达到最大值的点.

命题 5.6.5 若 $p<q$,则 f^* 为 g 的最小过剩主部.

证 按命题 5.6.4,f^* 为 g 的过剩主部. 设 $\Gamma\in\mathscr{E}$,并令 $r=\inf\{x\in\tilde{E}:x\in\Gamma\}$,则必有 $\tau_r=\inf\{n\in\mathrm{N}:\xi_n\in\Gamma\}$. 记

$$\mathscr{T}=\{\tau\in\mathscr{M}:\tau=\inf\{n\in\mathrm{N}:\xi_n\in\Gamma\},\Gamma\in\mathscr{E}\},$$

则 $$f^*(x)\leqslant\sup\{E_xg(\xi_\tau):\tau\in\mathscr{T}\}=\overline{S}(x)=S(x).$$

按定理 5.5.1,S 为 g 的最小过剩主部. 故 $f^*=S$.

命题 5.6.6 当 $p<q$ 时,$\tau_0^*=\tau_{r^*}$ 为$(0,\overline{S})$-最优停时,但不是$(0,S)$-最优停时.

证 命题 5.6.5 表明:$f^*(x)=E_xg(\xi_{\tau_{r^*}})=E_x(\xi_{\tau_0^*})$. 注意,按 $f_r(x)$之定义,当 $\xi_n\in\Gamma_{r^*}$时,$f^*(\xi_n)=\xi_n^+=g(\xi_n)$. 故 $\tau_0^*=\tau_{r^*}$. 这表明:τ_{r^*} 为 g 的$(0,\overline{S})$-最优停时.

按命题 5.6.2,$P_x(\overline{\lim_{n\to\infty}}g(\xi_n)=0)=1$. 注意,$r^*>0$. 因此,当 $x<r^*$ 时,$P_x(\tau_0^*=\infty)=P_x(\tau_{r^*}=\infty)=P_x(\xi_n^+\notin\Gamma_{r^*},n\in\mathrm{N})>0$. 这表明:$\tau_0^*$ 不是 g 的$(0,S)$-最优停时.

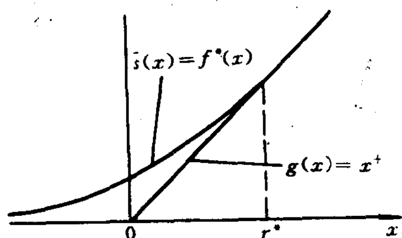


图 5.6-1 报酬函数 $\bar{S}(p<q)$

现将 $p<q$ 情形的结果如图 5.6-1 所示.

例 5.6.2 可观测过程具有两个吸收壁的随机游动.

设状态空间 $\bar{E}=\{0,1,\cdots,N\}$,状态 0 和 N 为吸收壁,即转移概率:$p(0,0)=1=p(N,N)$. 转移概率 $p(x,y)$定义为 $p(x,y)=$

$$P_x(\xi_1=y)=\begin{cases}\frac{1}{2}, 当 y=x+1 或 x-1 时.\\ 0, 其它.\end{cases} \quad x=1,2,\cdots,N-1. 定义$$

在 E 上的目标函数 g 是有界的,即 $|g(x)|<\infty(\forall\ x\in E)$.

假设 v 为 g 的最小过剩主部(即报酬函数),则 v 为如下方程的唯一解:

$$\begin{cases}v(x)=g(x)\vee Tv(x)\ (\forall\ x\in\bar{E}),\\ \overline{\lim_{n\to\infty}}v(\xi_n)=\overline{\lim_{n\to\infty}}g(\xi_n)(\bmod P_x)(\forall\ x\in\bar{E}).\end{cases}\tag{5.6-1}$$

(1) 当 $x=0$ 时,有

$$v(0)=g(0)\vee Tv(0)=g(0)\vee v(0).$$

按(5.6-1)的边界条件及 0 为吸收壁,得 $v(0)=g(0)$.

(2) 当 $x=N$ 时,有

$$v(N)=g(N)\vee Tv(N)=g(N)\vee v(N).$$

同(1)中的理由,得 $v(N)=g(N)$.

(3) 当 $1\leqslant x\leqslant N-1$ 时,有

$$v(x)\geqslant Tv(x)=\frac{1}{2}(v(x+1)+v(x-1)).$$

由此即知,v 为上凸函数,且

$$\frac{1}{2}(v(x+1)+v(x-1))$$

$$\geqslant \frac{1}{2}(g(x+1)+g(x-1)), v(x) \geqslant g(x). \quad (5.6\text{-}2)$$

因此,v 为满足条件(5.6-2)的最小上凸函数. 按定理 5.5.2,g 的 $(0,S)$-最优停时 τ_0 为

$$\tau_0 = \inf\{n \in \mathrm{N}: v(\xi_n) = g(\xi_n)\}, \inf\{\varnothing\} = \infty.$$

综上分析,$v(x)=S(x)$ $(\forall\ x\in E)$的几何构作方式是首先 v 在状态 0 和 N 处取值 $g(0)$和 $g(N)$. 然后找出 g 的凸凹段的分界点. 在 g 的凸部分,v 与 g 重合,在 g 的凹部分,v 为两个分界点的连线. 参看图 5.6-2 所示.

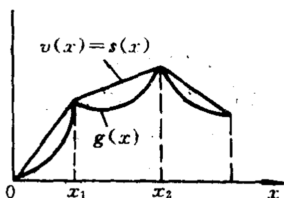


图 5.6-2 最小凸函数

例 5.6.3 观测过程具有两个反射壁的随机游动.

状态空间 E 和转移函数 $p(x,y)$同上例定义. 但不同点在于 $p(0,0)=0=p(N,N)$, $p(0,1)=1=p(N,N-1)$. g 的最小过剩主部 v 是满足下列要求的最小上凸函数

$$\begin{cases} v(x) \geqslant \frac{1}{2}(v(x+1)+v(x-1)), 1 \leqslant x \leqslant N-1; \\ v(0) \geqslant v(1); v(N) \geqslant v(N-1); v(x) \geqslant g(x), \forall\ x \in E. \end{cases} \quad (5.6\text{-}3)$$

由此即可推出

$$v(0) \geqslant v(1) \geqslant \cdots \geqslant v(N) \geqslant v(N-1)$$
$$\geqslant \cdots \geqslant v(1) \geqslant v(0).$$

故
$$\begin{cases} v(x) = v(0), v(x) \geqslant g(x)\ (\forall\ x \in E), \\ v(0) \geqslant \max\limits_{x\in E} g(x). \end{cases}$$

按定理 5.5.2,得

$$S(x) = v(x) = \max_{i\in E} g(i)\ (\forall\ x \in E);$$

$$\overline{\lim_{n\to\infty}} g(\xi_n) = \max_{i\in E} g(i) \ (\mathrm{mod} P_x) \ (\forall\ x \in E);$$

$$\tau_0 = \inf\{n \in \mathcal{N}: v(\xi_n) = g(\xi_n)\}$$

$$= \inf\{n \in \mathcal{N}: \max_{i\in E} g(i) = g(\xi_n)\}.$$

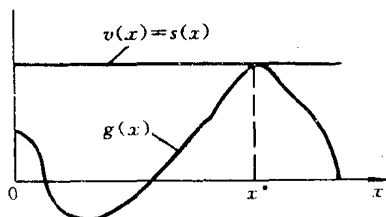


图 5.6-3　报酬函数 $\overline{S}$

此 τ_0 为 g 的 $(0,S)$-最优停时. 图 5.6-3 所示.

例 5.6.4　这里考虑例 5.6.2 和例 5.6.3 的混合情形.

假设 $p(0,0)=1, p(N,N-1)=1$(即 0 为吸收壁,N 为反射壁),类似的分析可知:v 是应满足下列要求的最小上凸函数:

$$\begin{cases} v(x) \geqslant \dfrac{1}{2}(v(x+1)+v(x-1)), 1 \leqslant x \leqslant N-1; \\ v(0) = g(0), v(N) \geqslant v(N-1); \\ v(x) \geqslant g(x) \ (\forall\ x \in E). \end{cases} \tag{5.6-4}$$

由条件(5.6-4)可推出下列结果:

(1) $v(N) \geqslant v(N-1) \geqslant \cdots \geqslant v(0)$. $v(N) \geqslant \max\limits_{x\in E} g(x)$. 由 v 的最小性,应取 $v(N) = \max\limits_{x\in E} g(x)$;

(2) 设 $x^* \in E$ 满足等式:$g(x^*) = \max\limits_{x\in E} g(x)$,且当 $x > x^*$ 时,$g(x) < g(x^*)$,则对 $\forall\ x \in E$,且 $x \leqslant x^*$,(5.6-4)可转化为

$$\begin{cases} v(x) \geqslant \dfrac{1}{2}(v(x+1)+v(x-1)), 1 \leqslant x \leqslant x^*-1; \\ v(0) = g(0), v(x^*) = g(x^*); \\ v(x) \geqslant g(x) \ (\forall\ x \in E). \end{cases} \tag{5.6-5}$$

这组方程的最小解 $v(x)(0\leqslant x\leqslant x^*)$ 和 $v(x)=g(x^*)(x\in E, x\geqslant x^*)$ 合起来即为所要求的解 $v(x)(x\in E)$. 图示如下.

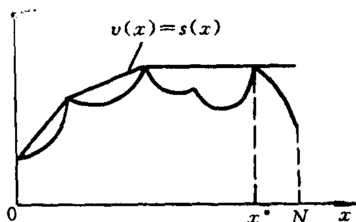


图 5.6-4　报酬函数 S

例 5.6.5　讨论所有状态都是反射壁的随机游动.

设 $E=\{0,1,2,\cdots\}$. 转移概率 $p(i,i+1)=1(\forall\ i\in E)$. 目标函数 $g:0\leqslant g(x)\uparrow K<\infty(\forall\ x\in E)$(严格上升). 显然, $g\in B^{-,+}$. 按定理 5.5-1, $\overline{S}(x)=S(x)=v(x)$ (g 的最小过剩主部).

相应这个模型的 v 是如下方程的最小解:

$$v(x)\geqslant v(x+1), v(x)\geqslant g(x)\ (\forall\ x\in E). \quad (5.6\text{-}6)$$

$$\Rightarrow v(x)\geqslant \sup_x g(x)=K.$$

由最小性得: $v(x)\equiv K(x\in E)$. 这时 $(0,\overline{S})$-最优停时 τ_0:

$$\tau_0=\inf\{n\in \mathrm{N}: v(\xi_n)=g(\xi_n)\}=\infty.$$

(ε,S)-最优停时 $\tau_\varepsilon(\varepsilon>0)$ 为

$$\tau_\varepsilon=\inf\{n\in \mathrm{N}: v(\xi_n)\leqslant g(\xi_n)+\varepsilon\}.$$

显然, τ_0 不是 $(0,S)$-最优停时. 参看图 5.6-5.

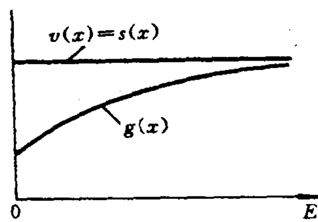


图 5.6-5　报酬函数 S

例 5.6.6　观测过程具有一个吸收壁的无限随机游动.

设状态空间 $E=\{0,1,2,\cdots\}$; $p(0,0)=1$(0 为吸收壁), $p(i,i+1)=\frac{1}{2}p(i,i-1)(i=1,2,\cdots)$. 目标函数 g 为

$$g(x)=1\vee x\ (\forall\ x\in E).$$

不难看出，g 的最小过剩主部 v 是如下方程的最小解：

$$\begin{cases}v(x)\geqslant\frac{1}{2}(v(x+1)+v(x-1)),x\geqslant 1;\\ v(0)=g(0),v(x)\geqslant g(x)\ (x\in E).\end{cases}\tag{5.6-7}$$

这个方程组的最小解 v 如图 5.6-6 所示.

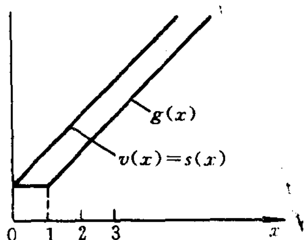


图 5.6-6　报酬函数 S

这种情形，有

$\tau_\varepsilon=\inf\{n\in N:v(\xi_n)\leqslant g(\xi_n)+\varepsilon\}\ (0<\varepsilon<1).$

$E_xg(\xi_{\tau_\varepsilon})=g(0)=1\ (\forall\ x\in E).$

$\tau_0=0.\ E_xg(\xi_{\tau_0})=g(x)\ (\forall\ x\in E).$

因此，当 $x\geqslant 2$ 时，有 $E_xg(\xi_{\tau_\varepsilon})<E_xg(\xi_{\tau_0})$. 由此可见，$\tau_\varepsilon$ 不是$(0.\overline{S})$-最优停时，出现这一现象的原因在于：g 不满足条件 A^+. 即

$$E_x\sup_n g^+(\xi_n)=\infty\ (\forall\ x\in E).$$

验证如下. 令 $\tilde{g}(x)=x^+$. $b\geqslant 1$ 为整数. $E^b=\{0,1,\cdots,b\}$. $\tilde{g}^{(b)}(x)=\tilde{g}(x)\wedge b$. $\tilde{g}^{(b)}$的最小过剩主部 $\tilde{v}^{(b)}(x)\equiv b$(按例 5.6.3). 按定理 5.5.1，有$\overline{\lim_{n\to\infty}}\tilde{g}^{(b)}(\xi_n)=\overline{\lim_{n\to\infty}}\tilde{v}^{(b)}(\xi_n)\equiv b$. 故

$$\overline{\lim_{n\to\infty}}\tilde{g}(\xi_n)=\infty\ (\mathrm{mod}P_x)\ (\forall\ x\in E).$$

如图 5.6-7 所示.

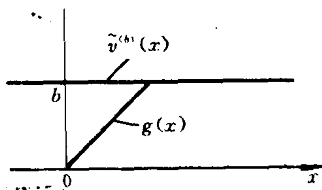


图 5.6-7　报酬函数 $\tilde{v}(b)$

例 5.6.7　设 $\eta=(\eta_n)_{n\geqslant 1}$ 为独立同分布的 r.v. 列，且 $P(\eta_n=0)=\frac{1}{2}=P(\eta_n=1)$. 令 $\xi_n=\prod_{k=0}^{n}(2\eta_k)$，其中 $\xi_0=2\eta_0$，且独立于 η. 状态空间 $E=\{0,2,2^2,\cdots\}$. 则 ξ 是以$(E,\mathscr{E})$为状态空间定义在$(\Omega,\mathscr{F},P_x)$上的齐

次 Markov 链($\forall\ x\in E$). 令 $\mathscr{F}_n=\sigma(\xi_0,\xi_1,\cdots,\xi_n)$,则 $\xi=(\xi_k)_{k\in\mathscr{N}}$ 为关于 P_x 的自生鞅($\forall\ x\in E$).

目标函数 $g(x)=-x^2(\forall\ x\in E)$.

(1) g 不满足条件 A^-. 事实上

$$E_x(\sup_n g^-(\xi_n))=E_x(\sup_n \xi_n^2)\geqslant E_x\xi_n^2$$

$$=2^{2(n+1)}\cdot x\cdot E_x(\prod_{k=1}^{n}\eta_k^2)=4\cdot 2^n x$$

$\Rightarrow$当 $x\neq 0$ 时,$E_x(\sup_n g^-(\xi_n))=\infty$,即 $g\notin A^-$.

(2) g 本身为过剩函数. 事实上,ξ^2 为下鞅,$-\xi^2$ 为上鞅. 这保证:$Tg(x)\leqslant g(x)(\forall\ x\in E)$. 故 g 为过剩函数.

(3) 目标函数 $g(x)=-x^2$ 使定理 5.5.1 不真. 事实上,假如定理 5.5.1 对 g 成立. 则应有 $S(x)=g(x)(\forall\ x\in E)$. 令

$$\tau^*=\inf\{n\in\mathrm{N}:\xi_n=0\}.$$

则 $P_x(\tau^*<\infty)=1(\forall\ x\in E)$. 为证实这一点,做如下计算. $P_x(\tau^*=0)=\begin{cases}1,x=0,\\0,x\neq 0.\end{cases}\Rightarrow P_x(\tau^*<\infty)=1$ 当 $x=0$ 时. 下面假定 $x\neq 0$,则对$\forall\ n\geqslant 1$,有

$$P_x(\tau^*=n)=P_x(\xi_i\neq 0,0\leqslant i\leqslant n-1;\xi_n=0)$$

$$=P_x(\eta_i=1,1\leqslant i\leqslant n-1;\eta_n=0)=\frac{1}{2^n}$$

$$\Rightarrow P_x(\tau^*<\infty)=\sum_{k=1}^{\infty}\frac{1}{2^n}=1.\ (\forall\ x\neq 0,x\in E).$$

令 $S^*(x)=E_xg(\xi_{\tau^*})=-E_x\xi_{\tau^*}^2=0$,则

$$S(x)=-x^2\leqslant 0=E_xg(\xi_{\tau^*})=S^*(x),$$

且

$$S(x)<S^*(x)\ (\forall\ x\neq 0,x\in E).$$

由 $\tau^*\in\mathscr{M}$ 可知:S 不是 g 的最优报酬函数. 这个矛盾表明:定理 5.5.1 不适用于这样的目标函数 g,究其原因在于:目标函数 $g\notin B^-$.

例 5.6.8 再次考查例 6.1 中的无限随机游动. 这里假定 $p=\frac{1}{2}=q$. 目标函数 $g(x)=x(\forall\ x\in E)$.

按例 5.6.6 中提供的方法可以证明：

$$E_x(\sup_n \xi_n^{\pm}) = \infty(\forall\ x \in E) \Rightarrow g \notin B^-, g \notin B^+.$$

又 $\xi=(\xi_n)_{n\in N}$关于 P_x 为自生鞅($\forall\ x\in E$). 因此, g 本身为过剩函数. 定义 $\tau^*=\infty$, 则因 $P_x(\overline{\lim}\xi_n=\infty)=1$ 而有 $\tau^*\in\overline{\mathscr{M}}_g$. 因此

$$\overline{S}(x) \geqslant E_x g(\xi_{\tau^*}) = \infty, \text{即} \overline{S}(x) = \infty\ (\forall\ x \in E).$$

如果定理 5.5.1 适用于此 g, 则应有 $\overline{S}(x)=g(x)(\forall\ x\in E)$. 但这不是事实. 究其原因在于: $g\notin B^-$.

注 1 例 5.6.1、5.6.8 表明: 当条件 A^- 不成立, 即 $g\notin B^-$ 时, g 的最小过剩主部 v 一般不是报酬函数 S. 但报酬函数 S 必属于 g 的过剩主部类. 因此, 为解决更一般的问题, 对 g 的过剩函数类需要加上某些限制.

§5.7 条件 a^- 下的报酬函数

在例 5.6.7 中, 由于 $g\notin B^-$ 而导致报酬函数 S 一般不是目标函数 g 的最小过剩主部. 但可验证: g 满足如下条件:

$$a^-: E_x g^-(\xi_\infty)=E_x((\overline{\lim_{n\to\infty}} g(\xi_n))^-)<\infty(\forall\ x\in E).$$

本节讨论在条件 a^- 下如何处理这种不一致的问题.

将目标函数 g 作如下修改:

$$G(x) = g(x) \vee E_x g(\xi_\infty)\ (\forall\ x \in E),$$

并以 G 为目标函数.

引理 5.7.1 设 $g\in B(a^-)$, 则

(i) 函数 $V(x)=\lim_{n\to\infty} Q^n G(x)(x\in E)$ 为 G 的最小过剩主部;

(ii) 报酬函数 $\overline{U}(x)=V(x)\ (x\in E)$, 其中

$$\overline{U}(x) = \sup\{E_x G(\xi_\tau): \tau \in \overline{\mathscr{M}}_g\}.$$

证 $g\in B(a^-)\Rightarrow E_x g(\xi_\infty)>-\infty\ (\forall\ x\in E)$

$\Rightarrow E_x G(\xi_1)\geqslant E_x E_{\xi_1} g(\xi_\infty)=E_x g(\xi_\infty)>-\infty.\Rightarrow g\in B_1.$

按P.10,V 为 G 的最小过剩主部.

令 $G_b(x)=G(x)\wedge b$,则 $G_b\in B^+(\forall\ b>0)$. 定义

$$\tilde{\tau}_\varepsilon=\inf\{n\in N: \tilde{V}_b(\xi_n)\leqslant G_b(\xi_n)+\varepsilon\},\inf\{\varnothing\}=\infty.$$

其中 $\tilde{V}_b$ 按P.14中的方式构作,则按(P.15)(i),有

$$P_x(\tilde{\tau}_\varepsilon<\infty)=1\ (\forall\ \varepsilon>0),\tilde{V}_b(x)\leqslant E_x\tilde{V}_b(\xi_{\tilde{\tau}_\varepsilon}).\ (\forall\ x\in E).$$

按P.14,$\tilde{V}_b$ 为 G_b 的过剩主部,假定 V_b 为 G_b 的最小过剩主部,则

$$V_b(x)\leqslant\tilde{V}_b(x)\leqslant E_x G_b(\xi_{\tilde{\tau}_\varepsilon})+\varepsilon\leqslant\overline{U}_b(x)+\varepsilon.$$

于是,$V_b\leqslant\overline{U}_b$. 考虑到 $g\in B(a^-)$,并利用单调收敛定理,得

$$V(x)=\lim_{b\to\infty}V_b(x)\leqslant\lim_{b\to\infty}\overline{U}_b(x)=\overline{U}(x).\ (\forall\ x\in E).$$

注意,$V(\xi_n)\geqslant G(\xi_n)\geqslant E_{\xi_n}g(\xi_\infty)=E_x(g(\xi_\infty)/\mathscr{F}_n)$,可得:

$$V^-(\xi_n)\leqslant E_x(g^-(\xi_\infty)/\mathscr{F}_n).\ (\forall\ n\in\mathrm{N}).$$

故 $(V^-(\xi_n))_{n\in\mathrm{N}}$ 一致可积($\because g^-(\xi_\infty)\in L^1$).

设 $\tau\in\overline{\mathscr{M}}$,$V_b(x)=V(x)\wedge b$,按P.5,$V_b$ 为过剩函数,且 $(V_b(\xi_n))_{n\in\mathrm{N}}$ 为一致可积上鞅. 按Doob鞅收敛定理,得 $\lim\limits_{n\to\infty}V_b(\xi_n)=V_b(\xi_\infty)(\mathrm{mod}P_x)(\forall\ x\in E)$,且

$$E_x(V_b(\xi_\infty)/\mathscr{F}_n)\leqslant V_b(\xi_n)\ (\mathrm{mod}P_x)\ (\forall\ n\in\mathrm{N},x\in E).$$

由鞅停时代替的不变性,有

$$E_x(V_b(\xi_\tau)/\mathscr{F}_\sigma)\leqslant V_b(\xi_\sigma)(\mathrm{mod}P_x)(\forall\ x\in E).$$

其中 σ、$\tau\in\overline{\mathscr{M}}$,$P_x(\sigma\leqslant\tau)=1(\forall\ x\in E)$. 由此便得

$$E_xV_b(\xi_\tau)\leqslant E_xV_b(\xi_\sigma)\leqslant V_b(x).\ (\forall\ x\in E).$$

按Fatou引理,得

$$E_xV(\xi_\tau)\leqslant V(x).\Rightarrow\overline{U}(x)\leqslant V(x)(\forall\ x\in E).$$

综上分析,得 $\overline{U}=V$.

定理 5.7.1 若 $g\in B(a^-)$,则

(i) $\overline{S}(x)=V(x)(\forall\ x\in E)$,其中 V 为 G 的最小过剩主部,其

中 $\overline{S}$ 为 g 的报酬函数.

(ii) $\overline{S}(x)=g(x)\vee E_x g(\xi_\infty)\vee T\overline{S}(x)(\forall\ x\in E)$.

证 按引理 5.7.1，若(i)成立，则(ii)自然成立. 下证(i). 由 $\overline{S}(x)\leqslant\overline{U}(x)$知：结论(i)$\Leftrightarrow\overline{U}(x)\leqslant\overline{S}(x)(\forall\ x\in E)$.

若 $\overline{S}(x)=\infty$，自然有 $\overline{U}(x)\leqslant\infty=\overline{S}(x)$. 下面讨论：$\overline{S}(x)<\infty$ 的情形.

令 $E^*=\{x\in E:\overline{S}(x)<\infty\}$，则 $g(\xi_\infty)=G(\xi_\infty)(\mathrm{mod}P_x)(\forall\ x\in E^*)$. 事实上，对$\forall\ x\in E^*$，有

$$
\begin{aligned}
G(\xi_\infty) &= \varlimsup_{n\to\infty} G(\xi_n) = \varlimsup_{n\to\infty}(g(\xi_n)\vee E_{\xi_n}g(\xi_\infty)) \\
&= \varlimsup_{n\to\infty}(g(\xi_n)\vee E_x(g(\xi_\infty)/\mathscr{F}_n)) \\
&\leqslant \varlimsup_{n\to\infty} g(\xi_n)\vee\varlimsup_{n\to\infty}E_x(g(\xi_\infty)/\mathscr{F}_n).
\end{aligned}
$$

由 $\overline{S}(x)<\infty$ 及 $g\in B(a^-)$ 可知：$E_x|g(\xi_\infty)|<\infty$($\because\ \tau=\infty$为停时). 按 Doob-Levy 定理(一致可积鞅收敛定理)，

$$G(\xi_\infty)\leqslant g(\xi_\infty)\vee E_x(g(\xi_\infty)/\mathscr{F}_\infty)=g(\xi_\infty)(\mathrm{mod}P_x).$$

按 G 之定义，有 $g(\xi_\infty)\leqslant G(\xi_\infty)$. 故 $g(\xi_\infty)=G(\xi_\infty)$.

现在，假如：$x\in E^*$，$\tau\in\overline{\mathscr{M}}$，且 $E_xG(\xi_\tau)$存在，则

$$E_xG(\xi_\tau)=E_xg(\xi_\sigma),$$

其中 $\sigma=\tau\cdot 1_D+\infty\cdot 1_{D^c}$，$D=\{\omega:\tau<\infty,E_{\xi_\tau}g(\xi_\infty)\leqslant g(\xi_\tau)\}$. 事实上，前一步已证明：$E_x(G(\xi_\tau)\cdot 1_{(\tau=\infty)})=E_x(g(\xi_\sigma)\cdot 1_{(\tau=\infty)})$当 $x\in E^*$时. 因此，剩下的问题是证明：

$$E_x(G(\xi_\tau)\cdot 1_{(\tau<\infty)})=E_x(g(\xi_\sigma)\cdot 1_{(\tau<\infty)}).$$

为此，令 $B=D^c\cap(\tau<\infty)$，则

$$
\begin{aligned}
E_x(G(\xi_\tau)\cdot 1_{(\tau<\infty)}) &= E_x(g(\xi_\tau)\cdot 1_D)+E_x(1_B\cdot E_{\xi_\tau}g(\xi_\infty)) \\
&= E_x(g(\xi_\tau)\cdot 1_D)+E_xE_x(1_B\cdot g(\xi_\infty)/\mathscr{F}_\tau) \\
&= E_x(g(\xi_\tau)1_D)+E_x(g(\xi_\infty)\cdot 1_B). \\
&= E_x(g(\xi_\sigma)\cdot 1_{(\tau<\infty)}).
\end{aligned}
$$

(注意：在第二个等式中用到一致可积鞅的停时代替不变性定理，

一致可积性来自：$E_x|g(\xi_\infty)|<\infty(\forall\ x\in E^*)$).

利用这步的结果，立即可得：对$\forall\ x\in E^*$，有

$$\overline{S}(x)=\sup\{E_xg(\xi_\sigma):\sigma\in\overline{\mathscr{M}}\}\geqslant\sup\{E_xG(\xi_\tau):\tau\in\overline{\mathscr{M}}\}=\overline{U}(x).$$

综上分析，得$\overline{S}(x)\geqslant\overline{U}(x)$.（$\forall\ x\in E$）.

注1 进一步考查例 5.6.7，在此例中：目标函数 g：

$$g(x)=-x^2(\forall\ x\in E).$$

在此例中，构作了一个有限停时 τ^*，使得 $E_xg(\xi_{\tau^*})=0$. 因此，$\overline{S}(x)\geqslant 0$ ($\forall\ x\in E$). 由 $g(x)\leqslant 0$ 知：$\overline{S}(x)\equiv 0$. 定义

$$\tau_n^*=\inf\{k\geqslant n:\xi_k=0\},\inf\{\varnothing\}=\infty.$$

则

$$P_x(\tau_n^*<\infty)=E_xE_{\xi_n}(1_{(\tau^*<\infty)})$$

$$=\sum_{y\in E}E_y1_{(\tau^*<\infty)}\cdot P_x(\xi_n=y)=1\ (\forall\ n\in\mathrm{N},x\in E)$$

$$\Rightarrow P_x(\bigcap_n(\tau_n^*<\infty))=1.\text{ 且 }\tau_n^*\uparrow\infty\ (\mathrm{mod}P_x).$$

注意：对$\forall\ \omega\in\bigcap_n(\tau_n^*<\infty)$，有 $\xi_{\tau_n^*}(\omega)=0$ ($\forall\ n\in\mathrm{N}$). 故

$$\overline{\lim_{n\to\infty}}g(\xi_n(\omega))=0\ (\mathrm{mod}P_x)\ (\forall\ x\in E)$$

$$\Rightarrow E_x(\overline{\lim_{n\to\infty}}g(\xi_n))=0\ (\forall\ x\in E)$$

$$\Rightarrow E_x(\overline{\lim_{n\to\infty}}g(\xi_n))^-=0,\text{即 } g\in B(a^-).$$

注意：$G(x)=g(x)\vee E_xg(\xi_\infty)$. 按上述定理，$\overline{U}(x)=\overline{S}(x)$. 因此，$\overline{U}(x)=0$. ($\forall\ x\in E$). 在例 5.6.7 中已证明：$\overline{S}$ 不是 g 的最小过剩主部. 因此，在条件 a^-满足的情形下，g 的报酬函数 $\overline{S}$ 可以不是 g 的最小过剩主部，但却为 G 的最小过剩主部. 这就得到了在这种情形下求 g 的报酬函数的一种处理方法.

§5.8 报酬函数的正则性

例 5.6.7 和例 5.6.8 表明：当目标函数 $g\notin B^-$时，g 的最小过剩主部不必是正则的，且其报酬函数 $\overline{S}$ 一般不是它的最小过剩主

部. 按P. 7, 当$g\in B^-$时, g的最小过剩主部是正则的, 且为它的报酬函数(定理5.5.1). 这些事实引导我们考查g的最小正则(过剩)主部问题.

定义5.8.1 设$\mathscr{U}\subset\overline{\mathscr{M}}$, $f\in B$. 若$E_xf(\xi_\tau)$有定义($\forall\ x\in E$, $\tau\in\mathscr{U}$), 且对任意σ、$\tau\in\mathscr{U}$, $P_x(\sigma\leqslant\tau)=1$($\forall\ x\in E$), 有

$$E_xf(\xi_\tau)\leqslant E_xf(\xi_\sigma)\ (\forall\ x\in E).$$

则称f是$\mathscr{U}$-正则的. "$\overline{\mathscr{M}}$-正则"简称"正则".

如果g、$f\in B$, 且f是$\mathscr{U}$-正则的; $g(x)\leqslant f(x)$ ($\forall\ x\in E$), 则称f为g的$\mathscr{U}$-正则主部.

注1 若$\tau=0$, $\sigma=1$均属于$\mathscr{U}$, 则$\mathscr{U}$-正则函数f必为过剩函数. 反之不真. 显然, 这里的正则性概念跟§5.1的一致.

定理5.8.1 设$g\in B^+$, 则

(i) $S(x)$是g的最小正则主部, 且

$$S(x)=\lim_{a\to-\infty}S_a(x)\ (\forall\ x\in E);$$

(ii) $\overline{S}(x)=S(x)$ ($\forall\ x\in E$),

其中$S_a(x)$为g_a的报酬函数,

$$g_a(x)=a\vee g(x);$$

(iii) $S(x)=g(x)\vee TS(x)$ ($\forall\ x\in E$);

(iv) $S(x)=\lim\limits_{b\to\infty}\lim\limits_{a\to-\infty}\lim\limits_{n\to\infty}Q^ng_a^b(x)$, 其中$g_a^b(x)=(a\vee g(x))\wedge b$.

证 第一步, 令$g_a(x)=a\vee g(x)$, S_a为g_a的报酬函数. $S_*(x)=\lim\limits_{a\to-\infty}S_a(x)$, 则$S_*\in B^+$, 且为$g$的正则主部, 此外,

$$S_*(x)=g(x)\vee TS_*(x), S_*(x)\geqslant\overline{S}(x)\geqslant S(x)\ (\forall\ x\in E).$$

事实上

$$S_*(x)=\lim_{a\to-\infty}\sup\{E_xg_a(\xi_\tau): \tau\in\overline{\mathscr{M}}\}\geqslant\lim_{a\to-\infty}\sup\{E_xg(\xi_\tau): \tau\in\overline{\mathscr{M}}\}$$

$$=\overline{S}(x)\geqslant S(x)\geqslant g(x)>-\infty, \Rightarrow S_*\in B.$$

注意: $S_a\in B^+$, 且$S_a\downarrow S_*$. 因此, $S_*\in B^+$. 从而, 可利用单调收敛定理得到

$$\lim_{a\to-\infty} TS_a(x) = T \lim_{a\to-\infty} S_a(x) = TS_*(x).$$

按定理 5.5.1，可推出

$$S_*(x) = \lim_{a\to-\infty} S_a(x) = \lim_{a\to-\infty} (g_a(x) \vee TS_a(x))$$
$$= g(x) \vee TS_*(x).$$

由 $S_a \in B^-$ 及 P.7，可得

$$E_x S_a(\xi_\tau) \leqslant E_x S_a(\xi_\sigma) \leqslant S_a(x) < \infty,$$

其中 $\sigma、\tau \in \overline{\mathscr{M}}, \sigma \leqslant \tau$. 利用单调收敛定理，由上式得

$$E_x S_*(\xi_\tau) \leqslant E_x S_*(\xi_\sigma) \leqslant S_*(x) \ (\forall\ \sigma、\tau \in \overline{\mathscr{M}}, \sigma \leqslant \tau).$$

综上分析，可得结论：S_* 为 g 的正则主部.

第二步：证明：$S_*(x) \leqslant S(x)$ $(\forall\ x \in E)$.

为此，对 $\forall\ \varepsilon > 0$，构作下列停时：

$$\sigma_\varepsilon^a = \inf\{n \in N: S_*(\xi_n) \leqslant g_a(\xi_n) + \varepsilon\}, \inf\{\varnothing\} = \infty;$$
$$\tau_\varepsilon^a = \inf\{n \in N: S_a(\xi_n) \leqslant g_a(\xi_n) + \varepsilon\}, \inf\{\varnothing\} = \infty;$$
$$\tau_\varepsilon^* = \inf\{n \in N: S_*(\xi_n) \leqslant g(\xi_n) + \varepsilon\}, \inf\{\varnothing\} = \infty;$$

则 $\sigma_\varepsilon^a \leqslant \tau_\varepsilon^a, \sigma_\varepsilon^a \leqslant \tau_\varepsilon^*, P_x(\tau_\varepsilon^a < \infty) = 1, P_x(\tau_\varepsilon^* < \infty) = 1, S_*(x) \leqslant E_x S_*(\xi_{\tau_\varepsilon^*})(\forall\ x \in E)$. 事实上，$g_a \in B^{-,+}$，因此，按定理 5.5.1 及 P.15，有

$$P_x(\tau_\varepsilon^a < \infty) = 1, S_*(x) = E_x S_a(\xi_{\tau_\varepsilon^a}).$$

由 $\sigma_\varepsilon^b \leqslant \tau_\varepsilon^a (\forall\ a \leqslant b \leqslant 0)$ 及 S_a 的正则性知

$$E_x S_a(\xi_{\tau_\varepsilon^a}) \leqslant E_x S_a(\xi_{\sigma_\varepsilon^b}) \leqslant S_a(x) \ (\forall\ a \leqslant b \leqslant 0)$$
$$\Rightarrow S_a(x) = E_x S_a(\xi_{\sigma_\varepsilon^b}) \ (\forall\ a \leqslant b \leqslant 0).$$

引用单调收敛定理，便可推出

$$S_*(x) = \lim_{a\to-\infty} S_a(x) = \lim_{a\to-\infty} E_x S_a(\xi_{\sigma_\varepsilon^b}) = E_x S_*(\xi_{\sigma_\varepsilon^b}) \ (\forall\ b \leqslant 0)$$
$$\Rightarrow -\infty < S_*(x) \leqslant E_x g_b(\xi_{\sigma_\varepsilon^b}) + \varepsilon$$
$$= E_x(g_b(\xi_{\sigma_\varepsilon^b}) \cdot 1_{\{g(\xi_{\sigma_\varepsilon^b}) \leqslant b\}}) + E_x(g_b(\xi_{\sigma_\varepsilon^b}) \cdot 1_{\{g(\xi_{\sigma_\varepsilon^b}) > b\}}) + \varepsilon$$
$$\leqslant b \cdot P_x(g(\xi_{\sigma_\varepsilon^b}) \leqslant b) + \alpha(\varepsilon, x),$$

其中 $$\alpha(\varepsilon,x)=E_x\sup_k g^+(\xi_k)+\varepsilon$$

$$\Rightarrow P_x(g(\xi_{\sigma_\varepsilon^b})\leqslant b)\leqslant\frac{1}{b}(S_*(x)-\alpha(\varepsilon,x))\ (\forall\ b<0)$$

$$\Rightarrow\lim_{b\to-\infty}P_x(g(\xi_{\sigma_\varepsilon^b})\leqslant b)=0\ (\because g\in B^+).$$

显然，$\tau_\varepsilon^*=\sigma_\varepsilon^b$ 于 $\{g(\xi_{\sigma_\varepsilon^b})>b\}$ 上. 故

$$\begin{aligned}P_x(\tau_\varepsilon^*<\infty)&\geqslant P_x(\tau_\varepsilon^*<\infty,g(\xi_{\sigma_\varepsilon^b})>b)\\&=P_x(\sigma_\varepsilon^b<\infty,g(\xi_{\sigma_\varepsilon^b})>b)\\&=1-P_x(g(\xi_{\sigma_\varepsilon^b})\leqslant b)\xrightarrow[b\to-\infty]{}1.\end{aligned}$$

现在，选一数列 $(b_i)_{i\geqslant1}$，$b_i\downarrow-\infty$，使得

$\lim\limits_{i\to\infty}1_{D_i}(\omega)=0$，其中 $D_i=\{g(\xi_{\sigma_\varepsilon^{b_i}})\leqslant b_i\}$，

则 $$S_*(x)=E_x(S_*(\xi_{\sigma_\varepsilon^{b_i}})(1_{D_i}+1_{D_i^c})).$$

由 $\tau_\varepsilon^*=\sigma_\varepsilon^{b_i}$ 于 D_i^c 上及 Fatou 引理可得

$$\begin{aligned}S_*(x)&\leqslant E_x(S_*(\xi_{\tau_\varepsilon^*})\cdot\overline{\lim_{i\to\infty}}1_{D_i^c})+E_x(\overline{\lim_{i\to\infty}}S_*(\xi_{\sigma_\varepsilon^{b_i}})\cdot1_{D_i})\\&=E_xS_*(\xi_{\tau_\varepsilon^*}).\end{aligned}$$

利用 τ_ε^* 和 $S_*(x)$ 的上述性质，便可得

$$\begin{aligned}&S_*(x)\leqslant E_xS_*(\xi_{\tau_\varepsilon^*})\leqslant E_xg(\xi_{\tau_\varepsilon^*})+\varepsilon\leqslant\bar{S}(x)+\varepsilon\\&\Rightarrow S_*(x)\leqslant\bar{S}(x)\ (\forall\ x\in E).\end{aligned}$$

第三步：证明：S_* 为 g 的最小正则主部.

事实上，若 f 为 g 的任意正则主部，则

$$\begin{aligned}&g(x)\leqslant f(x),E_xf(\xi_\tau)\leqslant E_xf(\xi_\sigma)\ (\forall\ \sigma、\tau\in\overline{\mathscr{M}},\sigma\leqslant\tau)\\&\Rightarrow f(x)\geqslant E_xf(\xi_\tau)\geqslant E_xg(\xi_\tau)\ (\forall\ \tau\in\overline{\mathscr{M}}).\\&\Rightarrow f(x)\geqslant\bar{S}(x)\geqslant S_*(x).\ (\text{按第二步}).\end{aligned}$$

第四步：归纳有关结论. 按第一、二步，得结论(i)、(ii)和(iii). 结论(iv)将是下面引理 5.8.2 的特例.

注. 2 按此定理，当 g 满足条件 A^+ 时，其报酬函数 $\bar{S}$ 为 g 的最小正则主部. 在这种情形下，g 的最小过剩主部一般不正则. 这

是因为算子 Q 与单步转移算子有关. 这种算子连续作用的结果不改变作用对象的正则性. 但是,从结论(iv)中可以看出:算子 Q 在报酬函数的结构中仍具有特别重要的意义.

引理 5.8.1 设 $g\in B^+$. 算子 G、函数 $\tilde{v}$ 及 φ 均按 P.14 中定义,则 $S(x)=\tilde{v}(x)(\forall\ x\in E)$.

证 令 $\varphi_a(x)=\varphi(x)\vee a, G_a\varphi_a(x)=g_a(x)\vee T\varphi_a(x)$, 则 $G_a^n\varphi_a(x)\geqslant G^n\varphi(x)$. 按 P.15(ii)及定理 5.5.1,有

$$S_a(x)=\tilde{v}_a(x)=\lim_{n\to\infty}G_a^n\varphi_a(x).$$

按定理 5.8.1(i),有 $S(x)=\lim\limits_{a\to-\infty}S_a(x)$. 利用单调收敛定理,可得

$$S(x)=\lim_{a\to-\infty}\lim_{n\to\infty}G_a^n\varphi_a(x)\geqslant\lim_{n\to\infty}G^n\varphi(x)=\tilde{v}(x).$$

即 $S(x)\geqslant\tilde{v}(x)(\forall\ x\in E)$. 剩下的问题就是证明反向不等式: $S(x)\leqslant\tilde{v}(x)(\forall\ x\in E)$.

按定理 5.8.1(iii),有 $S(x)=g(x)\vee TS(x)$. 于是, $G^nS(x)=S(x)(\forall\ n\in N)$. 按 S 与 φ 之定义,有 $S(x)\leqslant\varphi(x)$

$$\Rightarrow S(x)=\lim_{n\to\infty}G^nS(x)\leqslant\lim_{n\to\infty}G^n\varphi(x)=\tilde{v}(x)(\forall\ x\in E).$$

推论 5.8.1 若 $g\in B^+$,则

$$S(x)=\lim_{a\to-\infty}\lim_{n\to\infty}G_a^n\varphi_a(x)=\lim_{n\to\infty}\lim_{a\to-\infty}G_a^n\varphi_a(x)(\forall\ x\in E).$$

引理 5.8.2 (i)设 $g\in B$,则

$$S(x)=\lim_{b\to\infty}S^b(x),$$

$$S(x)=\lim_{b\to\infty}\lim_{a\to-\infty}\lim_{n\to\infty}Q^ng_a^b(x)(\forall\ x\in E).$$

(ii) 设 $g\in B^+$,则 $S(x)=\lim\limits_{a\to-\infty}S_a(x)$,且

$$S(x)=\lim_{a\to-\infty}\lim_{b\to\infty}\lim_{n\to\infty}\text{或}\lim_{a\to-\infty}\lim_{n\to\infty}\lim_{b\to\infty}$$

或 $$\lim_{b\to\infty}\lim_{a\to-\infty}\lim_{n\to\infty}Q^ng_a^b(x)(\forall\ x\in E).$$

证 设 $g\in B$. 显然, $S^b\uparrow$. 令 $S^*(x)=\lim\limits_{b\to\infty}S^b(x)$,则 $S(x)=S^*(x)(\forall\ x\in E)$. 为证此关系,只需证明: $S(x)\leqslant S^*(x)$.

设 $\tau\in\mathscr{M}_g$,即 $E_xg^-(\xi_\tau)<\infty$. 显然,

$$E_x(g^b(\xi_\tau))^- < \infty, g^b(\xi_\tau) \uparrow g(\xi_\tau).$$

因此,按单调收敛定理,有

$$\lim_{b\to\infty} E_x g^b(\xi_\tau) = E_x g(\xi_\tau).$$

又 $E_x g^b(\xi_\tau) \leqslant S^b(x) \leqslant S^*(x)$. 故

$$E_x g(\xi_\tau) \leqslant S^*(x)\ (\forall\, \tau \in \mathscr{M}_g, x \in E) \Rightarrow S(x) \leqslant S^*(x).$$

现在,按定理 5.8.1(i),有

$$S^b(x) = \lim_{a\to-\infty} S_a^b(x) = \lim_{a\to-\infty} \lim_{n\to\infty} Q^n g_a^b(x).$$

到此,我们证明了本引理中的第一个结论.

最后,按定理 5.5.1,有

$$S_a(x) = \lim_{b\to\infty} S_a^b(x) = \lim_{b\to\infty} \lim_{n\to\infty} Q^n g_a^b(x) = \lim_{n\to\infty} \lim_{b\to\infty} Q^n g_a^b(x).$$

按定理 5.8.1(i),有

$$S(x) = \lim_{a\to-\infty} S_a(x) = \lim_{a\to-\infty} \lim_{b\to\infty} \lim_{n\to\infty} Q^n g_a^b(x) = \lim_{a\to-\infty} \lim_{n\to\infty} \lim_{b\to\infty} Q^n g_a^b(x).$$

再利用本引理中的结论(i),即得(ii).

注 3 在三重极限公式中:当 $g \in B^+$ 时,只有 $\lim_{a\to-\infty}$ 与 $\lim_{n\to\infty}$ 的次序不可交换. 当 $g \in B$ 时,$\lim_{a\to-\infty}$ 和 $\lim_{b\to\infty}$ 均不能与 $\lim_{n\to\infty}$ 交换次序. 当 $g \in B^-$ 时,三重极限可以任意交换次序. 三重极限公式的重要意义在于:揭露了报酬函数与单步转移算子 Q 之间的关系.

定理 5.8.2 设 $g \in B^+$, S 为 g 的报酬函数,

$$\tau_\varepsilon = \inf\{n \in N; S(\xi_n) \leqslant g(\xi_n) + \varepsilon\}, \inf\{\varnothing\} = \infty\ (\varepsilon \geqslant 0).$$

则 (i) τ_ε 为 g 的 (ε, S)-最优停时 $(\varepsilon > 0)$;

(ii) τ_0 为 g 的 $(0, \overline{S})$-最优停时;

(iii) 若 $\tau_0 \in \mathscr{M}$,则 τ_0 为 g 的 $(0, S)$-最优停时;

(iv) 若 $\lim_{n\to\infty} g(\xi_n) = -\infty (\bmod P_x)(\forall\, x \in E)$,则 $\tau_0 \in \mathscr{M}$.

证 按定理 5.8.1(i),$S(x) = S_*(x)$. 从而,$\tau_\varepsilon = \tau_\varepsilon^*$, $P_x(\tau_\varepsilon < \infty) = 1$. 按引理 5.8.1 及 P.15,得

$$S(x) = \tilde{v}(x) \leqslant E_x\tilde{v}(\xi_{\tau_\varepsilon}) = E_xS(\xi_{\tau_\varepsilon}) \leqslant E_xg(\xi_{\tau_\varepsilon})+\varepsilon(\forall\, x \in E),$$

故 τ_ε 为 g 的 (ε, S)-最优停时 $(\varepsilon>0)$.

按 P.14，有

$$\overline{\lim_{n\to\infty}}\tilde{v}(\xi_n) = \overline{\lim_{n\to\infty}}g(\xi_n)(\mathrm{mod}P_x)(\forall\, x \in E).$$

按 P.12，有 $\tilde{v}(x)=E_x\tilde{v}(\xi_{n\wedge\tau_0})(\forall\, n\in\mathrm{N})$. 于是

$$\tilde{v}(x) = \lim_{a\to-\infty}\tilde{v}_a(x)$$

$$\begin{aligned}
\tilde{v}_a(x) &\leqslant \overline{\lim_{a\to-\infty}} E_x(\tilde{v}(\xi_{\tau_0})\cdot 1_{(\tau_0<n)}) \\
&\quad + \overline{\lim_{n\to\infty}}E_x(\tilde{v}_a(\xi_n)\cdot 1_{(\tau_0\geqslant n)}) \\
&\leqslant E_x(\tilde{v}_a(\xi_{\tau_0})\cdot 1_{(\tau_0<\infty)}) \\
&\quad + E_x\overline{\lim_{n\to\infty}}(\tilde{v}_a(\xi_n)\cdot 1_{(\tau_0\geqslant n)})(\text{按 Fatou 引理}) \\
&= E_x(\tilde{v}_a(\xi_{\tau_0})\cdot 1_{(\tau_0<\infty)}) \\
&\quad + E_x(\tilde{v}_a(\xi_\infty)\cdot 1_{(\tau_0=\infty)}) \\
&= E_x\tilde{v}_a(\xi_{\tau_0})
\end{aligned}$$

$$\Rightarrow \overline{S}(x) = \tilde{v}(x) \leqslant \lim_{a\to-\infty} E_x\tilde{v}_a(\xi_{\tau_0}) = E_x\tilde{v}(\xi_{\tau_0}) = E_x\overline{S}(\xi_{\tau_0})$$

$$\Rightarrow \overline{S}(x) = E_xg(\xi_{\tau_0}).$$

综上分析，结论(i)、(ii)真，而(iii)是(ii)的推论. 现在仅剩下证明结论(iv). 采用反证法. 若(iv)不成立，则$\exists\, x_0\in E$，$\ni P_{x_0}(\tau_0=\infty)>0$. 按条件$\overline{\lim_{n\to\infty}}g(\xi_n)=-\infty(\mathrm{mod}P_{x_0})$，有

$$\overline{S}(x_0) = E_{x_0}g(\xi_{\tau_0}) \leqslant E_{x_0}\sup_n g^+(\xi_n) + E_{x_0}(\overline{\lim}g(\xi_n)\cdot 1_{(\tau_0=\infty)})$$

$$= -\infty(\text{按 } g\in B^+)$$

$$\Rightarrow -\infty = \overline{S}(x_0) \geqslant g(x_0) > -\infty(\text{由 } g\in B). \text{ 这就得到矛盾.}$$

当 $g\in B^+$ 时，按定理 5.8.1，其报酬函数 S 是正则的. 下面的定理就是将此结论推广到 $g\in B$ 的情形.

定理 5.8.3 设 $g\in B$,则

(i) 报酬函数 S 为 g 的最小 $\mathscr{M}^g$-正则主部;

(ii) $S(x)=\overline{S}(x)(\forall\ x\in E)$;

(iii) $S(x)=\lim\limits_{b\to\infty}\lim\limits_{a\to-\infty}\lim\limits_{n\to\infty}Q^n g_a^b(x).(\forall\ x\in E)$;

(iv) 若 $g\in B_1$,则 $S(x)=g(x)\vee TS(x)(\forall\ x\in E)$.

证 设 $\tau\in\mathscr{M}_g$,即 $g\in B$,且 $E_x g^-(\xi_\tau)<\infty$. 按引理 5.3.2, $S(x)=\sup\{E_x g(\xi_\tau):\tau\in\mathscr{M}_g\}$. 设 S^b 为 $g^b=g\wedge b$ 的报酬函数,则易证:$S(x)=\lim\limits_{b\to\infty}S^b(x)$. 按定理 5.8.1(i),$S^b$ 为 g^b 的最小正则主部. 于是

$$E_x S^b(\xi_\tau)\leqslant E_x S^b(\xi_\sigma)(\forall\ \sigma,\tau\in\mathscr{M}_g,P_x(\sigma\leqslant\tau)=1)(\forall\ x\in E).$$

注意,$-g^-(\xi_\tau)\leqslant S^b(\xi_\tau)\uparrow S(\xi_\tau)$,$E_x g^-(\xi_\tau)<\infty(\forall\ \tau\in\mathscr{M}_g)$,则按单调收敛定理,可得 $E_x S(\xi_\tau)\leqslant E_x S(\xi_\sigma)$. 这表明:$S$ 是 $\mathscr{M}_g$-正则的.

设 f 为 g 的任意 $\mathscr{M}_g$-正则主部,则

$$E_x g(\xi_\tau)\leqslant E_x f(\xi_\tau)\leqslant f(x)(\forall\ \tau\in\mathscr{M}_g,x\in E)$$

$$\Rightarrow S(x)=\sup\{E_x g(\xi_\tau):\tau\in\mathscr{M}_g\}\leqslant f(x)\ (\forall\ x\in E).$$

综合起来得结论(i).

按定理 5.8.1(ii),$\overline{S^b}=S^b$. 按单调收敛定理,对 $\forall\ \tau\in\overline{\mathscr{M}}_g$,有

$$E_x g(\xi_\tau)=\lim_{b\to\infty}E_x g^b(\xi_\tau)\leqslant\lim_{b\to\infty}\overline{S}^b(x)=\lim_{b\to\infty}S^b(x)=S(x)$$

$$\Rightarrow S(x)\leqslant\overline{S}(x)\leqslant S(x),\text{即}\ \overline{S}(x)=S(x)(\forall\ x\in E).$$

这就是结论(ii).

按引理 5.8.2,

$$S^b(x)=\lim_{a\to-\infty}S_a^b(x)=\lim_{a\to-\infty}\lim_{n\to\infty}Q^n g_a^b(x)$$

$$\Rightarrow S(x)=\lim_{b\to\infty}\lim_{a\to-\infty}\lim_{n\to\infty}Q^n g_a^b(x).\Rightarrow\text{结论(iii)}.$$

设在证明最后的结论(iv). $g\in B_1$ 表明:$\tau\equiv1\in\mathscr{M}_g$. 又 $\sigma\equiv0\in\mathscr{M}_g$. 于是,按结论(i),有 $TS(x)=E_x S(\xi_1)\leqslant S(x)$. 由此知:$S(x)\geqslant g(x)\vee TS(x)$. 另一方面,按定理 5.8.1,有

$$S^b(x)=g^b(x)\vee TS^b(x)\Rightarrow S^b(x)\leqslant g(x)\vee TS(x)\ (\forall\ b>0)$$

$\Rightarrow S(x) \leqslant g(x) \vee TS(x)$. 故结论(iv) 成立.

定理 5.8.4 设 $g \in B^{-,+}$, $\tau_n^* = \inf\{0 \leqslant k \leqslant n: S_{n-k}(\xi_k) = g(\xi_k)\}$, $S_k(x) = Q^k g(x)$. 令 $\tau^* = \lim\limits_{n\to\infty} \tau_n^*$, 则

(i) τ^* 为 g 的$(0,\overline{S})$-最优停时;

(ii) 当 τ^* 有限时, τ^* 为 g 的$(0,S)$-最优停时;

(iii) 当 $g \in B^-$ 时, $\tau^* = \inf\{n \in N: S(\xi_n) = g(\xi_n)\}$, $\inf\{\varnothing\} = \infty$.

证 易证: $\tau_n^* \uparrow$. 因此, τ^* 存在, 且为停时. 如果(iii)成立, 则按定理 5.8.2, 结论(i)和(ii)自然成立.

为证(iii). 令 $\sigma = \inf\{n \in \mathrm{N}: S(\xi_n) = g(\xi_n)\}$, $\inf\{\varnothing\} = \infty$. 则结论(iii)的成立等价于证明: $\sigma = \tau^*\ (\mathrm{mod} P_x)\ (\forall\ x \in E)$. 设 $\omega \in (\sigma = n)$, 则 $g(\xi_i(\omega)) < S(\xi_i(\omega))\ (0 \leqslant i \leqslant n-1)$. 注意到: $S(x) = \lim\limits_{k\to\infty} S_k(x)$ ($\because g \in B^-$). 因此, 当 N' 足够大时, 有 $g(\xi_i(\omega)) < S_{N-i}(\xi_i(\omega))\ (0 \leqslant i \leqslant n-1)\ (\forall\ N \leqslant N')$. 从而

$$\tau^*(\omega) \geqslant n = \sigma(\omega) \Rightarrow \tau^* \geqslant \sigma \text{ 于} (\sigma < \infty) \text{ 上}.$$

现在, 设 $\omega \in (\sigma = \infty)$, 则 $S(\xi_i) > g(\xi_i)\ (\forall\ i \in \mathrm{N})$ 对此 ω 成立. 于是, 对$\forall\ n \in \mathrm{N}$, $\exists\ N' \in \mathrm{N}$, $\ni \tau_N^*(\omega) > n\ (\forall\ N \geqslant N')$. 这表明, $\tau^*(\omega) \geqslant n\ (\forall\ n \in \mathrm{N})$. 故 $\tau^*(\omega) = \infty$. 归纳起来, 得 $\sigma \leqslant \tau^*\ (\mathrm{mod} P_x)\ (\forall\ x \in E)$. 剩下的问题就是证明: $\sigma \geqslant \tau^*$. 为此, 让 $\omega \in (\tau_n^* = k)$, 则

$$g(\xi_i(\omega)) < S_{n-i}(\xi_i(\omega)) \leqslant S(\xi_i)\ (0 \leqslant i \leqslant k-1)$$

$$\Rightarrow \sigma(\omega) \geqslant \tau_n^*(\omega)\ (0 \leqslant k \leqslant n). \Rightarrow \sigma \geqslant \tau_n^*\ (\forall\ n \in \mathrm{N})$$

$$\Rightarrow \sigma \geqslant \tau^*\ (\mathrm{mod} P_x)\ (\forall\ x \in E).$$

定理 5.8.5 设 $g \in B^-$, σ^* 为 g 的$(0,S)$-最优停时. 若 $S(x) < \infty\ (\forall\ x \in E)$, 则 $\tau^* = \lim\limits_{n\to\infty} \tau_n^*$ 为 g 的$(0,S)$-最优停时, 且

$$P_x(\tau^* \leqslant \sigma^*) = 1\ (\forall\ x \in E).$$

证 σ^* 为 g 的$(0,S)$-最优停时表明: $\sigma^* \in \mathscr{M}_g$, 且

$$S(x) = E_x g(\xi_{\sigma^*}) \leqslant E_x S(\xi_{\sigma^*}).\ (\forall\ x \in E).$$

按定理 5.8.3, S 为 g 的最小 $\mathscr{M}_g$-正则主部, 因此, $E_x S(\xi_{\sigma^*}) \leqslant$

$S(x)$. 故

$$S(x)=E_xg(\xi_{\sigma^*})=E_xS(\xi_{\sigma^*})\ (\forall\ x\in E).$$

由此及条件 $S(x)<\infty(\forall\ x\in E)$ 可知：$E_x|S(\xi_{\sigma^*})|<\infty$. 于是

$$S(\xi_{\sigma^*})\geqslant g(\xi_{\sigma^*}),E_xS(\xi_{\sigma^*})=E_xg(\xi_{\sigma^*})$$
$$\Rightarrow S(\xi_{\sigma^*})=g(\xi_{\sigma^*})(\mathrm{mod}P_x)\ (\forall\ x\in E).$$

按定理 5.8.4(iii)，$\tau^*=\inf\{n\in\mathrm{N}:S(\xi_n)=g(\xi_n)\},\inf\{\varnothing\}=\infty$. 故 $\tau^*\leqslant\sigma^*$. 显然，$\tau^*\in\mathscr{M}_g(\because g\in B^-)$. 又 S 为 g 的最 $\mathscr{M}_g$-正则主部，故

$$E_xS(\xi_{\sigma^*})\leqslant E_xS(\xi_{\tau^*})\leqslant S(x)=E_xS(\xi_{\sigma^*})(\forall\ x\in E).$$

由此知：$S(x)=E_xS(\xi_{\tau^*})=E_xg(\xi_{\tau^*})$. $(\forall\ x\in E)$，即 τ^* 为 g 的 $(0,S)$-最优停时.

§5.9 递归方程

按§5.5、§5.8 中的分析，当 $g\in B_1$ 时，g 的报酬函数 S(g 的最小 $\mathscr{M}_g$-正则主部)和 g 的最小过剩主部 $S^*(x)=\lim\limits_{n\to\infty}Q^ng(x)$ 均满足如下方程：

$$\begin{cases} f(x)=g(x)\vee Tf(x); & (5.9\text{-}1)\\ \overline{\lim\limits_{n\to\infty}}f(\xi_n)=\overline{\lim\limits_{n\to\infty}}g(\xi_n)\ (\mathrm{mod}P_x) & (5.9\text{-}2)\end{cases}\quad (x\in E).$$

而且 S 与 S^* 一般不重合. 可见在 B_1 范围内，方程(5.9-1)和(5.9-2)的解一般不唯一. 但如果将 f 限制在 $\mathscr{M}_g$-正则函数类中，则方程(5.9-1)、(5.9-2)之解是唯一的. 本节讨论此方程之解的唯一性及不唯一时解的表现. 为此，首先对观测过程 ξ 作如下假设.

假设(B)：$P(n,x;\Gamma)=P_x(\xi_n\in\Gamma)$，有

$$\mu(n,x;\Gamma)=\frac{1}{n}\sum_{i=1}^{n}P(i,x;\Gamma)\ (\Gamma\in\mathscr{E},x\in E,n\in\mathrm{N}). \tag{5.9-3}$$

$$\mu(n,x;\cdot)\xrightarrow[n\to\infty]{w.}\mu(\cdot)\ (\forall\ x\in E)，(即弱收敛.) \tag{5.9-4}$$

即对∀ 有界 $\mathscr{E}$-可测实值函数 f，有

$$\int_E f(y)\mu(n,x;\mathrm{d}y) \xrightarrow[n\to\infty]{} \int_E f(y)\mu(\mathrm{d}y) \quad (\forall\ x\in E). \tag{5.9-5}$$

定理 5.9.1 设 $f_i\in B_1(i=1,2)$，且满足下列要求：

(i) f_i 满足方程(5.9-1). $(i=1,2)$；

(ii) $\sup\{|f_1(x)-f_2(x)|:x\in E\}<\infty$；

(iii) 在某个非空的 $\Lambda\in\mathscr{E}$ 上：$f_1(x)=f_2(x)$；

(iv) ξ 满足假设(B)；

(v) $\mu(E\backslash\Lambda)<1$，

则 $f_1(x)=f_2(x)$ $(\forall\ x\in E)$.

证 以下要利用如下初等不等式：

$$|a\vee b-a\vee c|=\begin{cases}|b-c|, & \text{当 } a\leqslant b,a\leqslant c \text{ 时},\\ b-a, & \text{当 } b\geqslant a\geqslant c \text{ 时},\\ c-a, & \text{当 } b\leqslant a\leqslant c \text{ 时},\\ 0, & \text{当 } a\geqslant b,a\geqslant c \text{ 时}.\end{cases}$$

$$\leqslant|b-c|.\ (\forall\ a、b、c\in R).$$

令 $r(x)=|f_1(x)-f_2(x)|$，则

$$r(x)=|g(x)\vee Tf_1(x)-g(x)\vee Tf_2(x)|$$
$$\leqslant|Tf_1(x)-Tf_2(x)|\leqslant Tr(x).$$

反复利用此不等式可得：$r(x)\leqslant T^n r(x)$，即

$$r(x)\leqslant\int_E r(y)P(n,x;\mathrm{d}y)$$

$$\Rightarrow r(x)\leqslant\int_E r(y)\mu(n,x;\mathrm{d}y)\ (\forall\ n\in\mathrm{N},x\in E).$$

按假设(B)，得

$$r(x)\leqslant\int_E r(y)\mu(\mathrm{d}y)\leqslant\sup\{r(y):y\in E\}$$
$$\cdot\mu(E\backslash\Lambda)\ (\forall\ x\in E).$$

$$\Rightarrow r(x)=0\ (\forall\ x\in E)\ (\text{按条件 v}).$$

例 9.1 设齐次 Markov 链 ξ 具有状态空间 $E=\{0,1\}$，转移阵为 $\mathscr{P}$，且满足要求：$|p_{00}+p_{11}-1|<1$. 令 $\lambda_1=1,\lambda_2=p_{00}+p_{11}-1$，则 $\mathscr{P}^n=\frac{1}{1-\lambda_2}(\mathscr{P}-\lambda_2 I)+\frac{\lambda_2^n}{\lambda_2-1}(\mathscr{P}-I)=(p_{ij}^{(n)})_{2\times 2}$.

假如 $p_{ij}>0(i,j\in E)$，则 $\lim\limits_{n\to\infty}p_{ij}^{(n)}=\begin{cases}\frac{1}{1-\lambda_2}p_{10}, & j=0,\\ \frac{1}{1-\lambda_2}p_{01}, & j=1.\end{cases}$ 于是，链 ξ 满足假设(B)，即

$$\mu(n,x;\{y\})=\frac{1}{n}\sum_{k=1}^{n}p_{xy}^{(k)}\xrightarrow[n\to\infty]{}\mu(\{y\}),$$

其中 $\mu(\{0\})=\frac{p_{10}}{1-\lambda_2},\mu(\{1\})=\frac{p_{01}}{1-\lambda_2}$.

考虑目标函数 g：$|g(x)|<\infty(\forall\ x\in E)$. 取 $\Lambda=\{0\}$. 则 $\mu(E\setminus\Lambda)=\frac{p_{01}}{1-\lambda_2}<1$. 假如 f_i 满足方程(5.9-1)，且 $f_1(0)=f_2(0)$，$|f_1(1)-f_2(1)|<\infty$，则按定理 5.9.1，有 $f_1(1)=f_2(1)$.

假如：$g(0)<g(1)$，则

$$g(x)\vee Tg(x)=\begin{cases}Tg(x) & x=0,\\ g(x), & x=1.\end{cases}$$

$$\Rightarrow S(x)=\begin{cases}\lim Q^n g(x)=\lim T^n g(x)=g(0)\cdot\mu(\{0\})\\ \qquad +g(1)\mu(\{1\}), & \text{当 } x=0 \text{ 时},\\ g(1). & \text{当 } x=1 \text{ 时}.\end{cases}$$

注意，S 满足方程(5.9-1). 因此，若 f 也满足方程(5.9-1)，且 $f(1)=g(1)$，则 $f=S$.

推论 5.9.1 若 ξ 满足假设(B)，且 $\mu(E)\leqslant 1$，f_i $i=1$、2，满足定理 5.9.1 中的要求(i)～(iv)，则 $f_1(x)=f_2(x)(\forall\ x\in E)$.

推论 5.9.2 设 ξ 满足条件 B，f、$g\in B$，且有界. f 满足方程(5.9-1)，且$\exists$ 非空 $\Lambda\in\mathscr{E}$，$\ni f(x)=g(x)(\forall\ x\in\Lambda)$，$\mu(E\setminus\Lambda)<1$，则

$f=S$，其中 S 为 g 的报酬函数.

证 g 有界表明：S 为 g 的最小过剩主部，按假设，f 为 g 的过剩主部. 因此 $S(x)\leqslant f(x)(\forall\ x\in E)$. 又 $S(x)\geqslant g(x)(\forall\ x\in E)$. 故 $S(x)=f(x)=g(x)$ 于 Λ 上，按定理 5.9.1，得 $S=f$.

注 1 假如 ξ 为一个不可约 Markov 链，且正常返，则 ξ 满足条件(B)，且 $\mu(E)=1,\mu(\{x\})>0(\forall\ x\in E)$.

定理 5.9.2 设 f_1,f_2 为方程(5.9-1)的 $\mathscr{E}$-可测解，且满足下列要求：

(i) $E_x(\sup\limits_n|f_1(\xi_n)-f_2(\xi_n)|)<\infty(\forall\ x\in E)$.

(ii) 对 $\forall\ \varepsilon>0$，$\exists$ 非空的 $\Lambda_\varepsilon\in\mathscr{E}$，便得

(a) $|f_1(x)-f_2(x)|<\varepsilon(\forall\ x\in\Lambda_\varepsilon)$；

(b) $P_x\{\xi_n\in\Lambda_\varepsilon$：有无穷多个 $n\in\mathrm{N}\}=1\ (\forall\ x\in E)$，

则 $f_1=f_2$.

证 令 $r(x)=|f_1(x)-f_2(x)|$. 在定理 5.9.1 的证明中已得到：$0\leqslant r(x)\leqslant T^n r(x)$. $(\forall\ x\in E,n\in\mathrm{N})$.

$$\Rightarrow E_x(r(\xi_{n+1})/\mathscr{F}_n)=E_{\xi_n}r(\xi_1)=Tr(\xi_n)\geqslant r(\xi_n)(\mathrm{mod}P_x),$$

即 $r(\xi.)=(r(\xi_n))_{n\in N}$ 关于 P_x 为非负下鞅. 条件(i)保证此下鞅 $r(\xi.)$ 一致可积. 从而，$\lim\limits_{n\to\infty}r(\xi_n)(\mathrm{mod}P_x)$ 存在. 条件(ii)(b)表明：对 $\forall\ \varepsilon>0$ 及 $\omega\in\Omega$，$\exists\ (n_k)_{k\geq1}$. 子 $\xi_{n_k}\in\Lambda_\varepsilon(\mathrm{mod}P_x)(\forall\ k\geqslant1)$. 条件(ii)(a)表明：

$$\lim_{n\to\infty}r(\xi_n)=\lim_{k\to\infty}r(\xi_{n_k})\leqslant\varepsilon\ (\mathrm{mod}P_x)$$

$$\Rightarrow\lim_{n\to\infty}r(\xi_n)=0\ (\mathrm{mod}P_x).$$

由 $r(\xi.)$ 的下鞅性及 Lebesque 有界收敛定理可推出：

$$0\leqslant r(x)\leqslant\lim_{k\to\infty}E_x r(\xi_{n_k})=E_x\lim_{k\to\infty}r(\xi_{n_k})=0\ (\forall\ x\in E).$$

注 2 上述定理不要求 ξ 满足假设 B.

推论 5.9.3 设 f_1、f_2 为方程(5.9-1)的 $\mathscr{E}$-可测解，且满足下列要求：

(i) $E_x(\sup\limits_n|f_1(\xi_n)-f_2(\xi_n)|)<\infty(\forall\ x\in E)$；

(ii) $\Lambda=\{x: |f_1(x)-f_2(x)|=0\}\neq\varnothing$，且

$P_x\{\xi_n\in\Lambda$：有无穷多个 $n\in\mathrm{N}\}=1(\forall\ x\in E)$，

则 $f_1=f_2$.

证 对$\forall\ \varepsilon>0$，有 $\Lambda\subset\Lambda_\varepsilon=\{x: |f_1(x)-f_2(x)|<\varepsilon\}$. 于是，按定理 5.9.2，结论成立.

推论 5.9.4 设 $g\in B$，f 为方程(5.9-1)的 $\mathscr{E}$-可测解，且满足下列要求：

(i) $E_x(\sup_n(|f(\xi_n)|+|g(\xi_n)|))<\infty(\forall\ x\in E)$，即 f、$g\in B^{-,+}$.

(ii) 存在非空 $\Lambda\in\mathscr{E}$，使得

(a) $f(x)=g(x)$ 于 $x\in\Lambda$ 上；

(b) $P_x(\xi_n\in\Lambda$：有无穷多个 $n\in N)=1(\forall\ x\in E)$.

则 $f=S$，其中 S 为 g 的报酬函数.

证 $g\in B^{-,+}$ 表明：S 为 g 的最小过剩主部. 按假设，f 为 g 的过剩主部. 因此，$g(x)\leqslant S(x)\leqslant f(x)(\forall\ x\in E)$. 按假设，此式中的等号成立于 $x\in\Lambda$ 上. 按推论 5.9.3，即可得结论.

定理 5.9.3 设 $g\in B^{-,+}$；$f\in B^+$，且满足方程(5.9-1)，则

$f=S\Leftrightarrow$(5.9-2)式成立.

即 g 的最小过剩主部. 为(5.9-1)与(5.9-2)的唯一解.

证 必要性就是 P.13. 下证充分性.

令 $\tau_\varepsilon=\inf\{n\in N: f(\xi_n)\leqslant g(\xi_n)+\varepsilon\},\inf\{\varnothing\}=\infty.$

按 P.12 及 $f\in B^+$，有

$$E_xf(\xi_{n\wedge\tau_\varepsilon})=f(x)\ (\forall\ n\in N,x\in E).$$

按条件(5.9-2)，对$\forall\ \varepsilon>0$，有

$$\begin{aligned}P_x(\tau_\varepsilon=\infty)&=P_x\{f(\xi_n)>g(\xi_n)+\varepsilon,n\in\mathrm{N}\}\\&=P_x\{\overline{\lim}f(\xi_n)>\overline{\lim}g(\xi_n)\}=0.\end{aligned}$$

即 $P_x(\tau_\varepsilon<\infty)=1.(\forall\ \varepsilon>0,x\in E)$.

由 $f\in B^+$ 及 Fatou 引理便可推出

$$f(x) \leqslant \varlimsup_{n\to\infty} E_x f(\xi_{n\wedge\tau_\varepsilon}) \leqslant E_x f(\xi_{\tau_\varepsilon})$$

$$\leqslant E_x g(\xi_{\tau_\varepsilon}) + \varepsilon \leqslant S(x) + \varepsilon. \Rightarrow f(x) \leqslant S(x) \ (\forall\ x \in E).$$

$g\in B^{-,+}$表明:S 为 g 的最小过剩主部,按 f 满足方程(5.9-1)知:f 为 g 的过剩主部,于是,$S(x)\leqslant f(x)$ $(\forall\ x\in E)$.综上分析,得 $f=S$.

注 3 上述定理中条件(5.9-2)的充分性(即唯一性),也可以看作是定理 5.9.2 的推论.不难验证,当条件(5.9-2)成立时,定理 5.9.2 中的条件均满足.这就可得:当 $g\in B^{-,+}$时,递归方程(5.9-1)和(5.9-2)在 B^+范围内有唯一解.然而,当 $g\notin B^{-,+}$时,我们已经知道方程(5.9-1)和(5.9-2)之解不必唯一.下面讨论方程(5.9-1)在这种情形下的解的一般结构.定义由 ξ 所诱导的尾 σ-代数:

$$\mathscr{F}_n^\infty = \sigma(\xi_n, \xi_{n+1}, \cdots), \mathscr{F}^\infty = \bigcap_n \mathscr{F}_n^\infty. \qquad (5.9\text{-}6)$$

定理 5.9.4 设 $g\in B^-$,η 为$\mathscr{F}^\infty$-可测的广义 r.v.,且 $E_x\eta^- < \infty$ $(\forall\ x\in E)$,$\bar{g}(x)=g(x)\vee E_x\eta$,$\bar{S}(\eta;x)=\sup\{E_x\bar{g}(\xi_\tau):\tau\in\overline{\mathscr{M}}\}$ $(\forall\ x\in E)$,则

$$\bar{S}(\eta;x) = \sup\{E_x((g(\xi_\tau)\cdot 1_{(\tau<\infty)} + \eta 1_{(\tau=\infty)}), \tau\in\overline{\mathscr{M}}\}.$$

且 $$\bar{S}(\eta;x) = g(x) \vee T\bar{S}(\eta;x) (\forall\ x \in E).$$

证 显然,$\bar{g}\in B^-$,按定理 5.5.1,有

$$\sup\{E_x\bar{g}(\xi_\tau):\tau \in \overline{\mathscr{M}}\} = \sup\{E_x\bar{g}(\xi_\tau):\tau \in \mathscr{M}\}(\forall\ x \in E).$$

$$\text{定义}\sigma(\omega) = \begin{cases}\tau(\omega), & \omega\in A = \{g(\xi_\tau)\geqslant E_{\xi_\tau}\eta\} \\ \infty, & \omega\in A^c\end{cases} \quad (\tau\in\mathscr{M}).$$

注意,$A\in\mathscr{F}_\tau$,则

$$E_x\bar{g}(\xi_\tau) = E_x(g(\xi_\tau)\cdot 1_A + (E_{\xi_\tau}\eta)\cdot 1_{A^c})$$

$$= E_x(g(\xi_\sigma)\cdot 1_{(\sigma<\infty)} + \eta\cdot 1_{(\sigma=\infty)}).$$

令 $\tilde{S}(x)=\sup\{E_x(g(\xi_\sigma)1_{(\sigma<\infty)}+\eta\cdot 1_{(\sigma=\infty)});\sigma\in\overline{\mathscr{M}}\}$,则

$$\bar{S}(\eta;x) = \sup\{E_x\bar{g}(\xi_\tau);\tau\in\mathscr{M}\} \leqslant \tilde{S}(x).$$

另一方面,$(E_x(\eta/\mathscr{F}_n))_{n\in N}$为广义鞅,且 $E_x\eta^- < \infty$.按广义 Doob

Levy 定理，有

$\overline{\lim} g(\xi_n) \geqslant \overline{\lim} E_{\xi_n}\eta = \overline{\lim} E_x(\eta/\mathscr{F}_n) \geqslant E_x(\eta/\mathscr{F}_\infty) = \eta.\ (\mathrm{mod}P_x).$

$$\Rightarrow \overline{S}(\eta;x) = \sup\{E_x(\overline{g}(\xi_\tau)) : \tau \in \overline{\mathscr{M}}\}$$

$$\geqslant \sup\{E_x(g(\xi_\tau)\cdot 1_{(\tau<\infty)} + \eta\cdot 1_{(\tau=\infty)}), \tau \in \overline{\mathscr{M}}\} = \overline{S}(x).$$

综合起来即可得定理中的第一个等式.

按定理 5.5.1，有

$$\overline{S}(\eta;x) = \overline{g}(x) \vee T\overline{S}(\eta;x) = g(x) \vee E_x\eta \vee T\overline{S}(\eta;x),$$

$$T\overline{S}(\eta;x) = E_x\overline{S}(\eta;\xi_1) \geqslant E_xE_{\xi_1}\eta = E_x\eta.$$

故定理中的第二个等式成立.

定理 5.9.5 设 $g\in B^{-,+}$；$f\in B^+$，且满足(5.9-1)式，则 f 有表示：

$$f(x) = \sup\{E_x(g(\xi_\tau)\cdot 1_{(\tau<\infty)} + \eta\cdot 1_{(\tau=\infty)}) : \tau \in \overline{\mathscr{M}}\}\ (\forall\ x\in E),$$

其中

$$\eta = \lim_{n\to\infty} f(\xi_n)\ (\mathrm{mod}P_x)(\forall\ x\in E).$$

证 按假设，有 $f\in B^{-,+}$. 从而，$(f(\xi_n))_{n\in N}$ 为一致可积上鞅. 于是，$\eta=\lim\limits_{n\to\infty} f(\xi_n)(\mathrm{mod}P_x)$ 存在且有限，且 $E_x|\eta|<\infty$.

令 $\overline{g}(x)=g(x)\vee E_x\eta$. 并注意：$E_x\eta\leqslant f(x)$(按定理 4.5.5)，则

$$\overline{g}(x) \vee Tf(x) = g(x) \vee Tf(x) = f(x)\ (\forall\ x\in E).$$

若
$$\overline{\lim} f(\xi_n) = \overline{\lim} g(\xi_n)(\mathrm{mod}P_x)(\forall\ x\in E) \quad (5.9\text{-}7)$$

则按定理 5.9.3(唯一性)，得 $f(x)=\overline{S}(\eta;x)(\forall\ x\in E)$. 按定理 5.9.4便得 f 的表示.

为证(5.9-7)式，仅需证明：$\overline{\lim} f(\xi_n)\leqslant\overline{\lim} g(\xi_n)$. 此式成立是因为

$$\overline{\lim}\,\overline{g}(\xi_n) \geqslant \overline{\lim} E_{\xi_n}\eta = \overline{\lim} E_x(\eta/\mathscr{F}_n) \geqslant \eta = \lim f(\xi_n).$$

注 4 通过上述讨论我们得到方程(5.9-1)的解族：

$$\{\overline{S}(\eta;x) : \eta \text{ 是 } \mathscr{F}^\infty\text{- 可测的广义 r. v. }, E_x\eta > -\infty\}(\forall\ x\in E).$$

若取 $\overline{\eta}=\overline{\lim} g(\xi_n)$，则当 $g\in B^-$ 时，g 的报酬函数 $\overline{S}$ 为：$\overline{S}(x)=\overline{S}(\overline{\eta};x)$. 此外，此解族的差别发生在$(\tau=\infty)$上. 因此，若最优停时

是有限的，这种差别就自动消失.

上一节的定理 5.8.3 表明：当 $g\in B$ 时，报酬函数 S 为 g 的最小 $\mathscr{M}_g$-正则主部. 这个事实使人们考虑是否能在 $\mathscr{M}_g$-正则函数范围内从方程(5.9-1)中找到所需要的解这样一个问题. 然而，验证 $\mathscr{M}_g$-正则性并不总是容易的. 这就刺激人们去寻找某些更容易验证的条件.

定理 5.9.6 设 $g\in B(a^-)$. 则满足方程(5.9-1)的函数 f 为 g 的 $\overline{\mathscr{M}}_g$-正则主部的充要条件是：f 满足如下方程：

$$f(x)=g(x)\vee E_xg(\xi_\infty)\vee Tf(x)(\forall\ x\in E),\qquad(5.9\text{-}8)$$

其中 $g(\xi_\infty)=\overline{\lim}g(\xi_n)$，$g^-(\xi_\infty)=(\overline{\lim}g(\xi_n))^-$.

证 先证必要性. $g\in B(a^-)$表明：$\tau\equiv\infty\in\overline{\mathscr{M}}_g$. 又 $\tau\equiv 0\in\overline{\mathscr{M}}_g$. 由 f 的正则性知：$E_xf(\xi_\infty)\leqslant f(x)(\forall\ x\in E)$. f 满足方程(5.9-1)表明：$g(\xi_\infty)\leqslant f(\xi_\infty)$. 于是

$$Tf(x)=E_xf(\xi_1)\geqslant E_xE_{\xi_1}g(\xi_\infty)=E_xg(\xi_\infty)(\forall\ x\in E).$$

$$\Rightarrow f(x)=g(x)\vee Tf(x)=g(x)\vee E_xg(\xi_\infty)\vee Tf(x)$$

$$(\forall\ x\in E).$$

往下证明充分性. 设 u 为方程(5.9-8)的解，则

$$u(x)\geqslant E_xg(\xi_\infty)=E_x(\overline{\lim}g(\xi_n))\geqslant -E_xg^-(\xi_\infty)=E_x\eta,$$

其中 $\eta=-g^-(\xi_\infty)$. 由此即知

$$u(\xi_n)\geqslant E_{\xi_n}\eta=E_x(\eta/\mathscr{F}_n)\ (\forall\ n\in\mathrm{N}).$$

按条件，$(u(\xi_n))_{n\in N}$关于 P_x 为受一个一致可积鞅下控的广义上鞅. 按停时代换的不变性，有

$$E_xg(\xi_\infty)\leqslant E_xu(\xi_\tau)\leqslant u(x)(\forall\ x\in E,\tau\in\overline{\mathscr{M}}_g).$$

$$E_xg(\xi_\infty)\leqslant E_xu(\xi_\infty)\leqslant E_xu(\xi_1)=Tu(x).\ (\forall\ x\in E).$$

这里利用了条件：$\tau\equiv\infty\in\overline{\mathscr{M}}_g$ 及 $g(\xi_\infty)=\overline{\lim}u(\xi_n)$. 故 u 满足方程(5.9-1)，且为 g 的 $\overline{\mathscr{M}}_g$-正则主部.

§5.10 截断最优停止规则

前面讨论的最优停时均与初始状态 x 无关. 在许多情况下，

如果将最优停时τ^*与x关联起来，则τ^*有可能对某些$x\in E$关于P_x是有限的或有界.本节讨论这类问题.

定义 5.10.1 设$\tau^*\in\mathscr{M}$为目标函数g的最优停时.若对点$x\in E$，存在有限整数$N(x)$，使得$P_x(\tau^*\leqslant N(x))=1$，则称$\tau^*$在点$x$处被截断.令$n(x)=\inf\{N(x):P_x(\tau^*\leqslant N(x))=1\}$，则称$n(x)$为$\tau^*$在点$x$处的截断数.如果存在有限整数$N$，使得对所有的$x\in E$，都有$P_x(\tau^*\leqslant N)=1$，则称$\tau^*$被截断.令$n=\inf\{N:P_x(\tau^*\leqslant N)=1,x\in E\}$，则称$n$为$\tau^*$的截断数.

本节总假定：目标函数$g\in B^-$.

引理 5.10.1 设$S_k(x)=Q^kg(x)$，$\alpha_k(x)=S_k(x)-TS_k(x)$，$\beta_k(x)=S_{k+1}(x)-S_k(x)$，$(\forall\ x\in E)$，则$\beta_k(\xi_{n-k})=0\Leftrightarrow\alpha_k(\xi_{n-k})\geqslant 0$.

证
$$\begin{aligned}\beta_k(x)=0&\Rightarrow g(x)\vee TS_k(x)-TS_k(x)\\&\quad+TS_k(x)-S_k(x)=0\\&\Rightarrow\alpha_k(x)=g(x)\vee TS_k(x)-TS_k(x)\geqslant 0.\end{aligned}$$
$$\begin{aligned}\alpha_k(x)\geqslant 0&\Rightarrow S_k(x)\geqslant TS_k(x)\Rightarrow\beta_k(x)\\&=g(x)\vee TS_k(x)-S_k(x)\leqslant 0\\&\Rightarrow S_{k+1}(x)-S_k(x)\leqslant 0\Rightarrow S_{k+1}(x)=S_k(x)(\because S_k(x)\uparrow)\\&\Rightarrow\beta_k(x)=0.\end{aligned}$$

引理 5.10.2 $\beta_{k+1}(x)\leqslant T\beta_k(x)$.$(\forall\ x\in E,k\in\mathrm{N})$.

证 利用如下初等不等式：
$$a\vee b-a\vee c\leqslant b-c\ (\forall\ b\geqslant c)$$
便可得
$$\begin{aligned}\beta_{k+1}(x)&=S_{k+2}(x)-S_{k+1}(x)\\&=g(x)\vee TS_{k+1}(x)-g(x)\vee TS_k(x)\\&\leqslant TS_{k+1}(x)-TS_k(x)=T\beta_k(x).\end{aligned}$$

定理 5.10.1 如果对于状态$x\in E$及某个确定的$k\geqslant 0$，存在有限整数$n_k=n_k(x)$，使得当$n\geqslant n_k$时，有

$$\beta_k(\xi_{n-k}) = 0(\bmod P_x), P_x(\beta_k(\xi_{n_k-k-1}) > 0) > 0.$$

则 $S_n(x)=S(x)$ ($\forall\ n\geqslant n_k$).

如果对于状态 $x\in E$ 及整数 k_1、k_2, $k_1\leqslant k_2$，存在有限整数 n_1、n_2，使得当 $n\geqslant n_i$ 时，有

$$\beta_{k_i}(\xi_{n-k_i}) = 0\ (\bmod P_x), P_x(\beta_k(\xi_{n_i-k_i-1}) > 0) > 0\ (i=1,2).$$

则 $n_2\leqslant n_1$.

证 按引理 5.10.2；$0\leqslant S_{n_k+1}(x)-S_{n_k}(x)\leqslant T\beta_{n_k-1}(x)\leqslant\cdots\leqslant T^{n_k-k}\beta_k(x)=E_x\beta_k(\xi_{n_k-k})=0$(按假设).

$$\Rightarrow\beta_{n_k}(x) = 0\Rightarrow\beta_n(x)\leqslant T^{n-k}\beta_k^{(x)} = E_x\beta_k(\xi_{n-k}) = 0.\ (\forall\ n\geqslant n_k).$$

$$\Rightarrow S_n(x) = S(x)\ (\forall\ n\geqslant n_k).$$ 这就得第一结论.

按引理 5.10.2，有

$$\begin{aligned}\beta_{k_2}(x)&\leqslant T^{k_2-k_1}\beta_{k_1}(x) = E_x\beta_{k_1}(\xi_{k_2-k_1})\\ &\Rightarrow\beta_{k_2}(\xi_{n-k_2})\leqslant E_{\xi_{n-k_2}}\beta_{k_1}(\xi_{k_2-k_1})\\ &=E_x(\beta_{k_1}(\xi_{k_2-k_1+n-k_2})/\mathscr{F}_{n-k_2})\\ &=E_x(\beta_{k_1}(\xi_{n-k_1})/\mathscr{F}_{n-k_2}) = 0\ (\forall\ n\geqslant n_1)\\ &\Rightarrow\beta_{k_2}(\xi_{n-k_2}) = 0(\forall\ n\geqslant n_1).\end{aligned}$$

已知：$\beta_{k_2}(\xi_{n-k_2})=0(\forall\ n\geqslant n_2)$. 故 $n_2\leqslant n_1$.

推论 5.10.1 如果对于状态 $x\in E$ 及某个 $k\geqslant 0$，存在有限整数 $n_k=n_k(x)$，使得 $\beta_k(\xi_{n-k})=0(\bmod P_x)(\forall\ n\geqslant n_k)$, $P_x(\beta_k(\xi_{n_k-k-1})>0)>0$. 则停时 $\tau_{n_k}^*$ 在点 x 处是最优的，即 $S(x)=E_xg(\xi_{\tau_{n_k}^*})$，其中

$$\tau_{n_k}^* = \inf\{0\leqslant m\leqslant n_k: S_{n_k-m}(\xi_m) = g(\xi_m)\}.$$

证 按定理 5.4.1, $S_{n_k}(x)=E_xg(\xi_{\tau_{n_k}^*})$. 按定理 5.10.1,

$$S(x) = S_n(x) = S_{n_k}(x)(\forall\ n\geqslant n_k).$$

推论 5.10.2 如果对某 $k\geqslant 0$，及 $\forall\ x\in E$，存在如定理 5.10.1 中所定义的有限整数 $n_k(x)$，使得 $n_k^*=\sup\{n_k(x): x\in E\}<\infty$，则停时 $\tau_{n_k^*}$ 是 g 的最优停时，即 $S(x)=E_xg(\xi_{\tau_{n_k^*}})(\forall\ x\in E)$.

证 由定理 5.4.1 及定理 5.10.1，立即可推出此结论.

引理 5.10.3 如果对于给定的状态 $x\in E$，在空间 $E^{n-m+1}(n-m\geqslant 0)$ 中，存在集 $\overset{n-m}{\underset{i=0}{\times}} A_i$，$\ni P_x(U_iU_{i+1})>0(0\leqslant i\leqslant n-m-1)$，其中 $U_{n-m}=\{\beta_m(\xi_{n-m})>0\}\cdot\{\xi_{n-m}\in A_{n-m}\}$；

$$U_i=\{Tg(\xi_i)\geqslant g(\xi_i)\}\cdot\{\xi_i\in A_i\}\ (0\leqslant i\leqslant n-m-1).$$

则 $\beta_i(\xi_{n-i})>0$ 于 U_{n-i} 上 $(m+1\leqslant i\leqslant n)$.

证 按假设，$P_x(U_iU_{i+1})>0(0\leqslant i\leqslant n-m-1)$. 于是，

$$\begin{aligned}1_{U_{n-m-1}}\beta_{m+1}(\xi_{n-m-1})&=(S_{m+2}(\xi_{n-m-1})-S_{m+1}(\xi_{n-m-1}))\cdot 1_{U_{n-m-1}}\\&=T(S_{m+1}-S_m)(\xi_{n-m-1})\cdot 1_{U_{n-m-1}}\\&=E_x(\beta_m(\xi_{n-m})/\mathscr{F}_{n-m-1})\cdot 1_{U_{n-m-1}}\\&\geqslant 1_{U_{n-m-1}}\cdot E_x(1_{U_{n-m-1}}1_{U_{n-m}}\beta_m(\xi_{n-m})/\mathscr{F}_{n-m-1}).\end{aligned}$$

按 U. 之定义，$\beta_m(\xi_{n-m})>0$ 于 U_{n-m} 上. 又 $P_x(U_{n-m-1}\cdot U_{n-m})>0$

$$\Rightarrow E_x(1_{U_{n-m-1}\cdot U_{n-m}}\beta_m(\xi_{n-m})/\mathscr{F}_{n-m-1})>0 \text{ 于 } U_{n-m-1} \text{ 上}.$$

$$\Rightarrow \beta_{m+1}(\xi_{n-m-1})>0 \text{ 于 } U_{n-m-1} \text{ 上}.$$

依此类推，即得 $\beta_i(\xi_{n-i})>0$ 于 U_{n-i} 上 $(m+1\leqslant i\leqslant n)$.

定理 5.10.2 如果对于状态 $x\in E$ 及某个 $k\geqslant 0$，存在有限整数 $n_k=n_k(x)$，使得

(i) $\beta_k(\xi_{n-k})=0(\mathrm{mod}P_x)(\forall\ n\geqslant n_k)$；

(ii) $P_x(\bigcap\limits_{i=0}^{n_k-k-2}\{Tg(\xi_i)\geqslant g(\xi_i)\}\cdot\{\beta_k(\xi_{n_k-k-1})>0\})>0.$

则 $S_{n_k-1}(x)<S_{n_k}(x)=S_{n_k+m}(x)=S(x)(\forall\ m\in\mathrm{N})$，且 n_k 为 τ^* 在点 x 处的截断数.

证 在引理 5.10.3 中，用 n_k-1 代 n，k 代 m，并取

$$A_i=\{Tg(x)\geqslant g(x)\}\ (0\leqslant i\leqslant n_k-1-k-1)$$

$$\{\xi_{n_k-1-k}\in A_{n_k-1-k}\}=\{\beta_k(\xi_{n_k-1-k})>0\}.$$

则按本定理之假设，引理 5.10.3 中的条件满足. 因此

$$\beta_i(\xi_{n_k-1-i})>0 \text{ 于 } \{\xi_{n_k-1-i}\in A_{n_k-1-i}\} \text{ 上 } (k+1\leqslant i\leqslant n_k-1)$$

$$\Rightarrow \beta_{n_k-1}(\xi_0)>0 \text{ 于 } \{\xi_0\in A_0\} \text{ 上}.$$

若 $x\notin A_0$，则因 $P_x(Tg(\xi_0)\geqslant g(\xi_0))=0$ 必导致与定理的假设矛盾. 因此，$x\in A_0$. 从而，$\beta_{n_k-1}(x)>0$，即 $S_{n_k-1}(x)<S_{n_k}(x)$. 按推论 5.10.1，便得定理中结论的第一部分. 按定义 5.10.1 及定理 5.10.1，n_k 为 τ^* 在 x 处的截断数. 这是因为

$$n_k(x)=\inf\{n\in \mathrm{N}:P_x(\beta_k(\xi_{j-k})=0)=1\ (j\geqslant n)\}.$$

§5.11　停时的随机化和充足类

在前面考虑的停时类中，我们固定 σ-代数流 $\mathscr{F}.=(\mathscr{F}_n)_{n\in N}$. 现在，要考虑 $\mathscr{F}.$ 的改变对问题的影响. 为此，记 $\mathscr{M}(\mathscr{F}.)=\mathscr{M}.\ S(x;\mathscr{F}.)=\sup\{E_xg(\xi_\tau):\tau\in\mathscr{M}(\mathscr{F}.)\}$. 一个值得注意的事实是：如果另有一个 σ-代数流 $\mathscr{F}^*.=(\mathscr{F}_n^*)_{n\in N}$不满足要求 $\mathscr{F}_n\subseteq\mathscr{F}_n^*\ (n\in N)$，则 ξ 就不一定保持它原有的特性. 就是因为这个事实，在改变 $\mathscr{F}.$ 时，必须满足这个包含关系. 现在的问题是：这种改变对 g 的报酬函数 S 有没有影响? 如果没有影响，那么如何利用这种任意性，选择更简单的 $\mathscr{F}.$ 和更小的停时类来得到 S?

首先作下列假设：

H-1 $\mathscr{F}.\subseteq\mathscr{F}^*.$，即 $\mathscr{F}_n\subseteq\mathscr{F}_n^*(\forall\ n\in \mathrm{N})$；

H-2 P_x^* 是 P_x 到 $\mathscr{F}_\infty^*=\sigma(\bigcup\limits_n\mathscr{F}_n^*)$上的扩张；

H-3 $\xi^*=(\xi_n)_{n\in N}$适应 $\mathscr{F}^*.$，且关于 P_x^* 为齐次 Markov 链 $(\forall\ x\in E)$.

定义 5.11.1　在假设 H-1 下，称$\mathscr{M}(\mathscr{F}^*.)$为关于 $\mathscr{F}.$ 的随机化停时类.

显然，$\mathscr{M}(\mathscr{F}^*.)\supseteq\mathscr{M}(\mathscr{F}.)$. $S^*(x)=\sup\{E_xg(\xi_\tau):\tau\in\mathscr{M}(\mathscr{F}^*.)\}\geqslant S(x)$.

定理 5.11.1　若 $g\in B$，则 $S^*=S$.

证　按定理 5.8.3，$S^*(x)=\lim\limits_{b\to\infty}\lim\limits_{a\to-\infty}\lim\limits_{n\to\infty}Q^ng_a^b(x)=S(x)$. 这个关系成立的原因在于：$Q^ng_a^b(x)$与 $\mathscr{F}.$ 无关.

注 1 按此定理，这种随机化并不影响报酬函数 S 的最终结果. 因此，从这一点出发，σ-代数流 $\mathscr{F}.$ 可选为：$\mathscr{F}_n=\sigma(\xi_0,\xi_1,\cdots,\xi_n)$ $(\forall\ n\in N)$，即 ξ 可以看作是一个自适应的齐次 Markov 链.

例 5.11.1 假如对某 $x\in E$，有 $S(x)=+\infty$，则在 $\mathscr{M}(\mathscr{F}.)$ 中可以没有最优停时存在. 但可以选择一个 σ-代数流 $\mathscr{F}^*.$（满足 H-1），于在 $\mathscr{M}(\mathscr{F}^*.)$ 中存在最优停时. 作法如下：设 $\tau_i\in\mathscr{M}(\mathscr{F}.)$，使得 $\infty=S(x)=\sup\limits_i E_x g(\xi_{\tau_i})$，$E_x g(\xi_{\tau_i})\geqslant 2^i\ (i\geqslant 1)$. 假定：$\rho=\rho(\omega)$ 为一个取值于 $\{1,2,\cdots\}$ 的 $\mathscr{F}$-可测的 r. v. ：$P(\rho=i)=2^{-i}$，且 ρ 与 $\mathscr{F}_\infty$ 独立.

第一步. 构作 σ-代数流 $\mathscr{F}^*.$：

$$\mathscr{F}_n^*=\sigma(\mathscr{F}_n;\{(\rho=i):1\leqslant i<\infty\})\supseteq\mathscr{F}_n.$$

第二步. 将 P_x 扩张到 $\mathscr{F}_\infty^*$ 上而成为 P_x^*：

$$P_x^*(A\cap(\rho=i))=P(\rho=i)\cdot P_x(A)\quad(\forall\ i\geqslant 1,A\in\mathscr{F}_\infty).$$

第三步. 定义关于 $\mathscr{F}^*.$ 的停时 τ^*：

$$\tau^*=\tau_i(\omega)\ \text{当}\ \omega\in(\rho=i)\ \text{时}(1\leqslant i<\infty).$$

这个停时 τ^* 是关于 P_x^* 是有限的，即 $P_x^*(\tau^*<\infty)=1$.

事实上，$(\tau^*=k)=\bigcup\limits_{i=1}^{\infty}(\rho=i)(\tau_i=k)\in\mathscr{F}_k^*\Rightarrow\tau^*$ 为 $\mathscr{F}^*.$-时.

$$P_x^*(\tau^*=k)=\sum_{i=1}^{\infty}P(\rho=i)\cdot P_x(\tau_i=k)$$

$$\Rightarrow P_x^*(\tau^*<\infty)=\sum_{k=0}^{\infty}P_x^*(\tau^*=k)$$

$$=\sum_{k=0}^{\infty}\sum_{i=1}^{\infty}P(\rho=i)P_x(\tau_i=k)=1.$$

第四步. 计算 $E_x^*g(\xi_{\tau^*})$：

$$E_x^*g(\xi_{\tau^*})=\sum_{i=1}^{\infty}E_x^*(g(\xi_{\tau_i})\cdot 1_{(\rho=i)})=\sum_{i=1}^{\infty}P(\rho=i)\cdot E_x^*g(\xi_{\tau_i})$$

$$=\sum_{i=1}^{\infty}2^{-i}\cdot E_x g(\xi_{\tau_i})=\infty.$$

$\Rightarrow S(x)=E_x^* g(\xi_{\tau^*})$. 故 $\tau^* \in \mathscr{M}(\mathscr{F}^*.)$ 为 g 的最优停时.

例 5.11.2 考虑停时的变分问题.

设 $\mathscr{M}_{c;f}^{x_0}(\mathscr{F}.)=\{\tau\in\mathscr{M}(\mathscr{F}.):E_{x_0}f(\xi_\tau)=C\}$. ($x_0\in E$, c 为常数)

$$\tilde{S}(x_0)=\sup\{E_{x_0}g(\xi_\tau):\tau\in\mathscr{M}_{c;f}^{x_0}(\mathscr{F}.)\}.$$

问题:设 f、$g\in B$. 如何随机化这个最优停时问题?

构作程序如下:

(1) 作出 τ_1、$\tau_2\in\mathscr{M}(\mathscr{F}.)$,使得

$$E_{x_0}f(\xi_{\tau_1})=a<c, E_{x_0}f(\xi_{\tau_2})=b>c.$$

(这里给出的 f 可以调整. 因此,总可以做到这一步).

(2) 构作 ρ 满足下列要求:

ρ 是 $\mathscr{F}$-可测的;ρ 独立于 $\mathscr{F}_\infty$;ρ 具有如下分布:

$$P(\rho=1)=\frac{b-c}{b-a}, P(\rho=2)=\frac{c-a}{b-a}, P(\rho\in\{1,2\})=1.$$

(3) 构作 $\mathscr{F}^*.$: $\mathscr{F}_n^*=\sigma(\mathscr{F}_n;\{(\rho=1),(\rho=2)\})$. ($n\in\mathbf{N}$).

(4) 构作 P_{x_0} 到 $\mathscr{F}_\infty^*$ 上的扩张 $P_{x_0}^*$:

$$P_{x_0}^*(A\cap(\rho=i))=P(\rho=i)\cdot P_{x_0}(A)\ (i=1,2,A\in\mathscr{F}_\infty).$$

(5) 构作 τ^*: $\tau^*=\tau_1\cdot 1_{(\rho=1)}+\tau_2 1_{(\rho=2)}$. $\tau^*\in\mathscr{M}_{c;f}^{x_0}(\mathscr{F}^*.)$.

事实上,$P_{x_0}^*(\tau^*<\infty)=P_{x_0}^*((\rho=i)\cdot(\tau_i<\infty),i=1,2)=P(\rho=1)+P(\rho=2)=1$.

$$E_{x_0}^*f(\xi_{\tau^*})=P(\rho=1)\cdot E_{x_0}f(\xi_{\tau_1})+P(\rho=2)\cdot E_{x_0}f(\xi_{\tau_2})$$

$$=\frac{b-c}{b-a}a+\frac{c-a}{b-a}b=c.\Rightarrow\tau^*\in\mathscr{M}_{c;f}^{x_0}(\mathscr{F}^*.).$$

(6) $\tilde{S}(x_0)=\sup\{E_{x_0}^*g(\xi_\tau):\tau\in\mathscr{M}_{c;f}^{x_0}(\mathscr{F}^*.)\}$.

(7) 设 $\tilde{S}_c(x_0)=\tilde{S}(x_0)$,且$\exists$ a、b,$a<c<b$,于 $\tilde{S}_a(x_0)$ 和 $\tilde{S}_b(x_0)$ 都容易计算,则 $\tilde{S}_c(x_0)=P(\rho=1)\cdot\tilde{S}_a(x_0)+P(\rho=2)\cdot\tilde{S}_b(x_0)$.

注 2 通过随机化方法总可以将问题转化成具有有限最优停时问题. 此方法的特点在于扩大停时类. 能否在更窄的停时类中寻

求最优停时呢？这个问题在前面就已经提出来了. 当 $g\in B^-$ 时，g 的报酬函数 $\overline{S}$ 可以在停时类 $\mathscr{M}(\mathscr{F}.)$ 中求出（$\mathscr{M}(\mathscr{F}.)\subseteq\overline{\mathscr{M}}(\mathscr{F}.)$）. 现在要考虑这个问题的一般情形.

定义 5.11.2 设 $\mathscr{F}.=(\mathscr{F}_n)_{n\in N}$. 若 $\mathscr{F}_n\subseteq\mathscr{F}_n(n\in\mathrm{N})$，且

$$\sup\{E_xg(\xi_\tau):\tau\in\mathscr{M}(\mathscr{F}.)\}=\sup\{E_xg(\xi_\tau):\tau\in\mathscr{M}(\mathscr{F}.)\}.$$

则称 $\mathscr{F}.$ 为充足 σ-代数流.

定理 5.11.2 设 $\xi=(\xi_n)_{n\in N}$ 为适应于 $\mathscr{F}.$ 且关于 P_x 为齐次 Markov 链. $g\in B^+$. $\widetilde{\mathscr{F}}.=(\widetilde{\mathscr{F}}_n)_{n\in N}$ 为 σ-代数流. 若

(i) $g(\xi_n)$ 是 $\widetilde{\mathscr{F}}_n$-可测的（$\forall\ n\in\mathrm{N}$）

(ii) $E_x(z/\mathscr{F}_n)=E_x(z/\widetilde{\mathscr{F}}_n)(\mathrm{mod}P_x)(x\in E,n\in\mathrm{N})$，

其中 z 为 $\widetilde{\mathscr{F}}_{n+1}$-可测的 r. v. 且 $E_x|z|<\infty(\forall\ x\in E)$.

证 按定理 5.8.1，$S(x)=\lim\limits_{b\to\infty}\lim\limits_{a\to-\infty}\lim\limits_{n\to\infty}Q^ng_a^b(x)$.

按条件(i)，$S(\xi_n)$ 是 $\widetilde{\mathscr{F}}_n$-可测的. 令

$$\tau_\varepsilon=\inf\{n\in\mathrm{N}:S(\xi_n)\leqslant g(\xi_n)+\varepsilon\}(\varepsilon>0).$$

则按定理 5.8.2，τ_ε 是 g 的 (ε,S)-最优停时，且 $\tau_\varepsilon\in\mathscr{M}(\widetilde{\mathscr{F}}.)$

$$\Rightarrow S(x)\leqslant\sup\{E_xg(\xi_\tau):\tau\in\mathscr{M}(\widetilde{\mathscr{F}}.)\}+\varepsilon$$

$$\Rightarrow S(x)\leqslant\sup\{E_xg(\xi_\tau):\tau\in\mathscr{M}(\widetilde{\mathscr{F}}.)\}.$$

反向不等式显然成立. 因为 $\mathscr{M}(\widetilde{\mathscr{F}}.)\subset\mathscr{M}(\mathscr{F}.)$. 故定义 5.11.1 中的条件满足，即 $\widetilde{\mathscr{F}}.$ 为充分 σ-代数流.

第六章　假设检验中的最优决策

假设检验在统计学中是一个相当重要的问题. 我们在这一章作为应用专题只考虑两个假设检验问题. 着重讨论鞅和最优停时在 π-Bayes 最优决策问题中的表现形式,并介绍有关结果.

§6.1　一般决策问题

设 ξ 是$(\Omega,\mathscr{F})\to(E,\mathscr{E})$的随机元. 在决策问题中它被看作一个观测对象. 记

$$\mathscr{F}^{\xi}=\sigma(\xi)=\{A:A=\xi^{-1}(B),B\in\mathscr{E}\}.$$

Θ表示某个参数集. $\mathscr{P}=(P_\theta)_{\theta\in\Theta}$为可测空间$(\Omega,\mathscr{F}^{\xi})$上的概率测度族. 显然,对$\forall\ \theta\in\Theta,P_\theta$ 为 ξ 的一个分布. 例如,$E=\{0,1,2,\cdots\}$,$\Theta=(\theta)_{\theta>0}$,则

$$P_\theta(\xi=x)=\frac{\theta^x}{x!}e^{-\theta}(\theta\in\Theta,x\in E).$$

定义 ξ 的一个概率分布族. 在决策问题中,假定 ξ 的分布属于族 $\mathscr{P}$,但不知 ξ 的具体分布:或者全不知或者部分不知. 统计推断就是要求通过观测数据得到 ξ 的分布信息. 而决策问题就是通过观测数据对 ξ 的分布为某 P_θ 之假设作出判决并分析这种判决所引起的风险或损失. 数学描述如下,设$(D,\mathscr{D})$为决策可测空间,$L(\theta,d)$ $(\theta\in\Theta,d\in D)$表示 ξ 的分布为 P_θ 的决策所引起的损失,称为损失函数. 假设,对$\forall\ \theta\in\Theta$,$\exists$ 唯一一个正确决策 $d\in D$,$\exists\ L(\theta,d)=0$. 对每个观测值 $x\in E$,都可作出一个相应的决策 $d_x\in D$,其损失为 $L(\theta,d_x)$. 显然,按概率分布 P_θ 考虑平均损失比较合理. 基于这种考虑,自然要求作出的决策所导致的平均损失最小. 设

$$\Delta=\{d:d\text{ 是}(\Omega,\mathscr{F}^{\xi})\to(D,\mathscr{D})\text{ 的随机元}\}.$$

按可测函数的表现定理，对$\forall\ d\in\Delta$，有$d=d(\xi)$，称$d\in\Delta$为决策规则（或决策函数）. 令

$$R(\theta,d)=E_\theta L(\theta,d(\xi))\ (\theta\in\Theta,d\in\Delta),$$

其中E_θ表示关于P_θ的期望算子，称$R(\theta,d)$为风险函数.

下面介绍决策问题的几种类型.

问题Ⅰ 设$\theta\in\Theta$，要求找出一个决策函数d^*，使得

$$E_\theta L(\theta,d^*)\leqslant E_\theta L(\theta,d)\ (\forall\ d\in\Delta).$$

问题Ⅱ 设$L(\theta,d)=(L_1(\theta,d),L_2(\theta,d))$，即二维损失函数矢量. 令

$$\Delta_\alpha=\{d\in\Delta:E_\theta L_1(\theta,d(\xi))\leqslant\alpha(\forall\ \theta\in\Theta)\},$$

其中$\alpha>0$，要求找到一个$d^*\in\Delta_\alpha$，使得

$$E_\theta L_2(\theta,d^*(\xi))\leqslant E_\theta L_2(\theta,d(\xi))\ (\forall\ \theta\in\Theta,d\in\Delta_\alpha).$$

问题Ⅲ 要求寻找一个$\theta^*\in\Theta,d^*\in\Delta$，使得

$$R(\theta^*,d^*)=\min_{d\in\Delta}\max_{\theta\in\Theta}R(\theta,d).$$

问题Ⅳ 对事先指定的$c>0$，要求找到$\theta\in\Theta$和$d\in\Delta$，使得

$$R(\theta,d)\leqslant c.$$

这里c表示可接受的风险界.

问题Ⅴ 设$(\Theta,\mathscr{U})$为参数可测空间，π为定义在此空间上的$r.v.\ \theta$的一个先验概率分布. π-Bayes最优决策问题是：要求找到一个$d^*\in\Delta$，使得

$$E^\pi R(\theta,d^*)=\min_{d\in\Delta}E^\pi R(\theta,d).$$

其中E^π表示关于分布π的期望算子.

定义6.1.1 设G为$(E,\mathscr{E})$到自身的变换群.

若(i)对$\forall\ g\in G$，g保持$\mathscr{P}$不变，即当P_θ为$g(\xi)$的一个分布时，$P_\theta\in\mathscr{P}$.

(ii)设$P_\theta,P_{\theta'}$分别为ξ、$g(\xi)$的分布，则由关系

$$P_\theta(\xi\in B)=P_{\theta'}(g(\xi)\in B)\ (\forall\ B\in\mathscr{E}).$$

定义一个变换$\theta'=\bar{g}(\theta):\Theta\to\Theta$，且此变换$\bar{g}$是1-1对应的.

(iii)对$\forall\ g\in G$，$\exists\ g^*=h(g):D\to D$，使得h为一个同态，即

$$h(g_1\circ g_2)=h(g_1)\circ h(g_2)\ (\forall\ g_1、g_2\in G).$$

且损失函数在此变换下不变，即

$$L(\bar{g}(\theta),g^*(d))=L(\theta,d)\ (\forall\ \theta\in\Theta,d\in D),$$

则称此统计决策问题在此变换下不变.

定义 6.1.2 若决策规则$d\in\Delta$满足要求

$$E_\theta L(\theta',d(\xi))\geqslant E_\theta L(\theta,d(\xi))\ (\forall\ \theta,\theta'\in\Theta),$$

则称d是无偏的.

定义 6.1.3 若$d,d'\in\Delta$，且

$$R(\theta,d')\leqslant R(\theta,d)\ (\forall\ \theta\in\Theta).$$

又 $$\exists\ \theta\in\Theta,\ni R(\theta_0,d')<R(\theta_0,d),$$

则说d受d'控制.称d是不可允许的.如果Δ中没有任何元素控制d，则称d是可允许的.如果Δ中的子类Δ'具有性质：对$\forall\ d\in\Delta'$，$\exists\ d'\in\Delta'$，$\ni d'$可控制d，则称Δ'为一个完备决策类.如果在完备类Δ'中不含任何完备子类，则称Δ'为最小完备决策类.

注 1 Δ中的元素凡受控的均可删去而不影响最优决策问题.对于π-Bayes 最优决策，所有的 Bayes 解d^*及它们的极限全体构成一个最小完备类.

定义 6.1.4 设τ是$(\Omega,\mathscr{F}^\xi)\to(\mathrm{N},\mathscr{N})$的可测映射，其中$\mathrm{N}=\{0,1,2,\cdots\}$，$\mathscr{N}$为N中子集之全体.若对$\forall\ B\in\mathscr{E}$，条件概率$P_\theta(\xi\in B/\tau)$与$\theta\in\Theta$无关，则称$\tau$为关于$\mathscr{P}$(或$\Theta$)的一个充分统计.

定理 5.1.1 设E为一个离散空间.τ为关于$\mathscr{P}$的一个充分统计的充要条件是，对$\forall\ \theta\in\Theta$，存在因子分解.

$$p_\theta(x)=g_\theta(\tau(x))\cdot h(x)\ (\forall\ x\in E),$$

其中$p_\theta(x)=P_\theta(\xi=x)$，$h(x)(x\in E)$与$\theta$无关，$g_\theta(n)\ (\theta\in\Theta,n\in\mathscr{N})$与$x$无关.

证 必要性.设τ为$\mathscr{P}$的一个充分统计，按可测函数的表现定理，有$\tau=t(\xi)$.因此，对$\forall\ \theta\in\Theta$，有

$$p_\theta(x)=E_\theta E_\theta(1_{\{\xi=x\}}/\tau)$$

$$= E_\theta(1_{\{\tau(\xi)=\tau(x)\}} E_\theta(1_{\{\xi=x\}/\tau(\xi)}))$$
$$= E_\theta(1_{\{\tau(\xi)=\tau(x)\}} P_\theta(\xi = x/\tau(\xi) = \tau(x)))$$
$$= P_\theta(\tau = \tau(x)) P_\theta(\xi = x/\tau = \tau(x))$$
$$= g_\theta(\tau(x)) \cdot h(x),$$

其中 $g_\theta(\tau(x))=P_\theta(\tau=\tau(x)), h(x)=P_\theta(\xi=x/\tau=\tau(x))$.
按假设, $h(x)$ 与 θ 无关.

充分性 设 $p_\theta(x)=g_\theta(\tau(x)) \cdot h(x)$, 则

$$P_\theta(\xi= x/\tau = \tau(x)) = \frac{P_\theta(\xi = x, \tau(\xi) = \tau(x))}{P_\theta(\tau = \tau(x))}$$
$$= \frac{P_\theta(\xi = x)}{P_\theta(\tau = \tau(x))} = \frac{g_\theta(\tau(x)) \cdot h(x)}{P_\theta(\tau = \tau(x))}.$$

$$P_\theta(\tau = \tau(x)) = \sum_{y\in E} P_\theta(\tau = \tau(x)/\xi = y) \cdot P_\theta(\xi = y)$$
$$= \sum_{y\in E} P_\theta(\xi = y) 1_{B(x)}(y)$$
$$= \sum_{y\in E} g_\theta(\tau(y)) h(y) 1_{B(x)}(y)$$
$$= g_\theta(\tau(x)) \sum_{y\in E} h(y) \cdot 1_{B(x)}(y).$$

其中 $B(x)=\{y:\tau(y)=\tau(x)\}$. 综合起来,得

$$P_\theta(\xi = x|\tau = \tau(x)) = h(x)\Big(\sum_{y\in E} h(y) 1_{B(x)}(y)\Big)^{-1}.$$

由此即知, $P_\theta(\xi=x|\tau)$ 与 $\theta\in\Theta$ 无关.

§6.2 假设检验问题

这一节介绍统计决策中一类很重要的实际问题,即通过统计资料检验某假设 H 的正确性. 让"假设 H"$\triangle$"ξ 具有概率分布 P_θ". 现在要求对接受还是拒绝"假设 H"作出决策, d_0 表示接受, d_1 表示拒绝,即决策空间 D 由二个元素组成,将 $\mathscr{P}$, Θ 分解成两个不同的部分:

$$\mathscr{P}=\mathscr{P}_0\cup\mathscr{P}_1,\Theta=\Theta_0\cup\Theta_1.$$

当 $\theta\in\Theta_0$ 或 $P_\theta\in\mathscr{P}_0$ 时，接受“假设 H”. 当 $\theta\in\Theta_1$ 或 $P_\theta\in\mathscr{P}_1$ 时，拒绝“假设 H”

定义 6.2.1 记 $S_i(d)=\{x\in E:d(x)=d_i\}$ $(d\in\Delta,i=0,1)$，若 $\theta\in\Theta_0$，则称 $P_\theta(\xi\in S_1(d))=P_\theta(d(\xi)=d_1)$ 为第一类偏差（即假设 H 正确而被拒绝的概率）. 若 $\theta\in\Theta_1$，则称 $P_\theta(\xi\in S_0(d))=P_\theta(d(\xi)=d_0)$ 为第二类偏差（即假设 H 不正确而被接受的概率）.

如果能将这两类偏差同时控制到最小程度则能得到相对比较理想的决策，但一般很难实现. 因此，在假设检验中一般放弃这种考虑而提出下列决策问题.

问题 I 设

$$\Delta_\alpha=\{d\in\Delta:P_\theta(d(\xi)=d_1)\leqslant\alpha\ (\forall\ \theta\in\Theta_0)\},$$

其中 α 为一个可接受的正常数，现在要求找到一个决策函数 $d^*\in\Delta_\alpha$，使得

$$P_\theta(d^*(\xi)=d_0)=\min_{d\in\Delta_\alpha}P_\theta(d(\xi)=d_0)\ (\forall\ \theta\in\Theta_1)$$

或 $$P_\theta(d^*(\xi)=d_1)=\max_{d\in\Delta_\alpha}P_\theta(d(\xi)=d_1)\ (\forall\ \theta\in\Theta_1).$$

问题 II 设

$$L_1(\theta,d_1)=1_{\Theta_0}(\theta),\ L_1(\theta,d_0)=0\ (\forall\ \theta\in\Theta);$$

$$L_2(\theta,d_1)=0\ (\forall\ \theta\in\Theta),L_2(\theta,d_0)=1_{\Theta_1}(\theta);$$

$$\Delta_\alpha=\{d\in\Delta:E_\theta L_1(\theta,d(\xi))\leqslant\alpha\ (\forall\ \theta\in\Theta)\};$$

$$\beta(\theta,d)\triangleq P_\theta(d(\xi)=d_0)=E_\theta L_2(\theta,d(\xi))\ (\theta\in\Theta_1),$$

则最优决策问题是寻找 $d^*\in\Delta\alpha$，使得

$$\beta(\theta,d^*)=\min_{d\in\Delta_\alpha}\beta(\theta,d)\ (\forall\ \theta\in\Theta_1)$$

或 $$\bar{\beta}(\theta,d^*)=\min_{d\in\Delta_\alpha}\bar{\beta}(\theta,d)\ (\forall\ \theta\in\Theta),$$

其中 $$\bar{\beta}(\theta,d)=E_\theta L_2(\theta,d(\xi))\ (\forall\ \theta\in\Theta).$$

注 1 易见，问题 I、II 等价. 仅区别于表述形式.

问题 III 记 $\varphi(\xi)=1_{\{d(\xi)=d_0\}}$，则

$$\varphi:(\Omega,\mathscr{F}^{\xi})\to([0,1],\mathscr{B}\cap[0,1]).$$

$$\beta(\theta,d)=P_{\theta}(d(\xi)=d_0)=E_{\theta}\varphi(\xi)\ (\theta\in\Theta_1);$$

$$\varphi(x)=0\text{ 或 }1\ (\forall\ x\in E).$$

按照$\beta(\theta,d)$的解释，其值越大，放弃这个θ的根据就越充分. 因此，可以称$\beta(\theta,d)$为功效函数，且$\beta(\theta,d)=E_{\theta}\varphi(\xi)\ (\theta\in\Theta_1)$；称$\varphi$为检验函数. 这时，最优决策，问题(Ⅱ)可转化成如下形式. 设

$$\Phi=\{\varphi:\varphi\text{为}(\Omega,\mathscr{F}^{\xi})\to([0,1],\mathscr{B}\cap[0,1])\text{ 的可测映射}\}.$$

$$\Phi_{\alpha}=\{\varphi\in\Phi:E_{\theta}\varphi(\xi)\leqslant\alpha\ (\forall\ \theta\in\Theta_0)\}\ (\alpha>0).$$

要求找到一个$\varphi^{*}\in\Phi_{\alpha}$，使得，对$\forall\ \theta\in\Theta_1$，有

$$\beta_{\varphi^{*}}(\theta)=\max_{\varphi\in\Phi_{\alpha}}\beta_{\varphi}(\theta),$$

其中
$$\beta_{\varphi}(\theta)=E_{\theta}\varphi(\xi).$$

问题Ⅳ 设E为一个离散空间. $\mathscr{P}=(P_i)_{i=0,1}$.

$$\text{“假设 }H\text{”}\triangleq\text{“}\xi\text{ 服从分布 }P_0\text{”};$$

$$\text{“假设 }K\text{”}\triangleq\text{“}\xi\text{ 服从分布 }P_1\text{”};$$

令$p_i(x)=P_i(\xi=x)\ (i=0,1,x\in E)$，则最优决策问题表述如下，要求确定判断域$S^{*}\subseteq E$，使得

$$\rho(S^{*})=\max\Big\{\rho(S)=\sum_{x\in S}p_1(x):S\subseteq E,\sum_{x\in S}p_0(x)\leqslant\alpha\Big\}.$$

定理 6.2.1 (Neyman-Pearson 定理) 设$\xi:(\Omega,\mathscr{F})\to(E,\mathscr{E})$的随机元，其中$E$为离散空间. P_0,P_1为$(\Omega,\mathscr{F}^{\xi})$上的两个不同的概率测度. $p_i(x)=P_i(\xi=x)\ (i=0,1,x\in E)$、定义$(E,\mathscr{E})$上的分布：$\widetilde{P}_i(B)=P_i(\xi\in B)\ (\forall\ B\in\mathscr{E})\ (i=0,1)$. 设$\mu$为$(E,\mathscr{E})$上的有限测度，且$\widetilde{P}_i\ll\mu\ (i=0,1,)$.（例如，可选$\mu=\widetilde{P}_0+\widetilde{P}_1$）. $\nu_i(x)=\dfrac{\mathrm{d}\widetilde{P}_i}{\mathrm{d}\mu}(x)\ (x\in E,i=0、1)$为$\widetilde{P}_i$关于$\mu$的 R.-N. 导数. 则

(i)在“检验H反对K”的决策问题中，

$$\exists\ \varphi^{*}\in\Phi\text{ 及 }r_0\in\overline{\mathbb{R}}_{+}=[0,\infty],\ni$$

(a) $E_0\varphi^{*}(\xi)=\alpha.$

(b) $\varphi^*(x)=\begin{cases}1, & \text{当 } \nu_1(x)>r_0\nu_0(x) \text{时},\\ 0, & \text{当 } \nu_1(x)<r_0\nu_0(x) \text{时}.\end{cases}$ $(x\in E)$.

(ii)若 $\varphi^*\in\Phi$，且满足(a)和(b)，则 φ^* 必为在水平 α 下"检验 H 反对 K"的最强检验函数，即对$\forall\ \varphi\in\Phi_\alpha$，

有 $E_1\varphi(\xi)\leqslant E_1\varphi^*(\xi)$.

(iii)若 φ^* 为在水平 α 下"检验 H 反对 K"的最强检验函数，则 $\exists\ r_0\in\overline{\mathbb{R}}_+$，∋要求(b)对 φ^* (modμ)成立，φ^* 也满足要求(a)，除非存在一个检验函数，其值小于 α，功效为 1.

证 若 $\alpha=0$ 或 1，则选取 $r_0=+\infty$或 0，即知，定理成立. 以下假定 $0<\alpha<1$.

先证(i). 设 $\alpha(c)=P_0(\nu_1(\xi)>c\nu_0(\xi))$. 由 $\widetilde{P}_0\ll\mu$ 知

$$\begin{aligned}\alpha(c)&=P_0(\nu_1(\xi)>c\nu_0(\xi),\nu_0(\xi)>0)\\&=P_0(r(\xi)>c,\nu_0(\xi)>0)\\&=P_0(r(\xi)>c)\\&=1-P_0(r(\xi)\leqslant c).\end{aligned}$$

其中 $P_0(r(\xi)\leqslant c)$ 为 $r(\xi)$ 在 P_0 下的分布函数. 因此，$\alpha(c)$对于 $c\in\overline{\mathbb{R}}_+$是右连续的，且单调不增. 显然，有 $\alpha(0)=0$，$\alpha(+\infty)=1$，$\alpha(c-0)-\alpha(c)=P_0(r(\xi)=c)$. 因此，对给定的 $0<\alpha<1$，方程

$$\alpha(c)\leqslant\alpha\leqslant\alpha(c-0)$$

总有解 $c_0\in\overline{\mathbb{R}}_{+0}=(0,\infty)$. 定义

$$\tilde{\varphi}(x)=\begin{cases}1, & \text{当 } \nu_1(x)>c_0\nu_0(x) \text{ 时},\\ 0, & \text{当 } \nu_1(x)<c_0\nu_0(x) \text{ 时}.\end{cases}$$

若 $\alpha(c_0-0)=\alpha(c)$，则定义 $\tilde{\varphi}(x)=1$ 于$\{\nu_1(x)=c_0\nu_0(x)\}$上. 若 $\alpha(c_0)<\alpha(c_0-0)$，则定义

$$\tilde{\varphi}(x)=\frac{\alpha-\alpha(c_0)}{\alpha(c_0-0)-\alpha(c_0)}.\ \text{于}\{\nu_1(x)=c_0\nu_0(x)\}\ \text{上}.$$

取 $r_0=c_0$，则 $\tilde{\varphi}$ 满足要求(b)，$\tilde{\varphi}$ 也满足要求(a). 事实上

$$E_0\tilde{\varphi}(\xi)=P_0(r(\xi)>c_0)+\frac{\alpha-\alpha(c_0)}{\alpha(c_0-0)-\alpha(c_0)}P_0(r(\xi)=c_0)$$

$$= \alpha(c_0) + \alpha - \alpha(c_0) = \alpha.$$

故 $\tilde{\varphi}$满足(a). 令 $\varphi^* = \tilde{\varphi}$. 则 φ^* 即为所求.

次证(ii) 设 $\varphi^* \in \Phi$, 且$\exists$ $r_0 \in \overline{\mathbb{R}}_+$, 使要求(a)和(b)成立. 让 $\varphi \in \Phi_\alpha$, 令

$$S^+ = \{x \in E: \varphi^*(x) - \varphi(x) > 0\},$$
$$S^- = \{x \in E: \varphi^*(x) - \varphi(x) < 0\}.$$

从 $S^\pm$的构成中可以得出:

$$x \in S^+ \Rightarrow \varphi^*(x) > 0 \Rightarrow \nu_1(x) \geqslant r_0\nu_0(x);$$
$$x \in S^- \Rightarrow \varphi^*(x) < 1 \Rightarrow \nu_1(x) \leqslant r_0\nu_0(x).$$

于是, 下列关系成立:

$$\int_E (\varphi^* - \varphi)(\nu_1 - r_0\nu_0)\mathrm{d}\mu = \int_{S^+ \cup S^-} (\varphi^* - \varphi)(\nu_1 - r_0\nu_0)\mathrm{d}\mu \geqslant 0.$$

$$E_1(\varphi^*(\xi) - \varphi(\xi)) \geqslant r_0 E_0(\varphi^*(\xi) - \varphi(\xi))$$
$$= r_0(\alpha - E_0\varphi(\xi)) \geqslant 0.$$

故
$$E_1\varphi(\xi) \leqslant E_1\varphi^*(\xi).$$

这表明, φ^* 是最强检验函数.

最后证(iii), 设 φ^* 为水平 α 下的最强检验函数. 按结论(i)和(ii), 有

$$E_1\varphi^*(\xi) = E_1\tilde{\varphi}(\xi), E_0\varphi^*(\xi) \leqslant \alpha = E_0\tilde{\varphi}(\xi).$$

令 $B = \{x \in E: (\tilde{\varphi}(x) - \varphi^*(x))(\nu_1(x) - r_0\nu_0(x)) \neq 0\}$,

则当 $x \in B$, 且 $\nu_1(x) > r_0\nu_0(x)$时, $\tilde{\varphi}(x) - \varphi^*(x) = 1 - \varphi^*(x) > 0$.

当 $x \in B$, 且 $\nu_1(x) < r_0\nu_0(x)$时, $\tilde{\varphi}(x) - \varphi^*(x) = -\varphi^*(x) < 0$.

于是, $(\tilde{\varphi}(x) - \varphi^*(x))(\nu_1(x) - r_0\nu_0(x)) > 0$ $(\forall\ x \in B)$.

假如 $\mu(B) > 0$, 则必有

$$\int_B (\tilde{\varphi} - \varphi^*)(\nu_1 - r_0\nu_0)\mathrm{d}\mu > 0.$$

然而

$$\int_B (\tilde{\varphi} - \varphi^*)(\nu_1 - r_0\nu_0)\mathrm{d}\mu = \int_E (\tilde{\varphi} - \varphi^*)(\nu_1 - r_0\nu_0)\mathrm{d}\mu$$

$$= E_1(\widetilde{\varphi}(\xi) - \varphi^*(\xi)) - r_0E_0(\widetilde{\varphi}(\xi) - \varphi^*(\xi))$$
$$= - r_0(\alpha - E_0\varphi^*(\xi)) \leqslant 0.$$

这就导致矛盾. 故 $\mu(B)=0$,即

$$(\widetilde{\varphi}(x) - \varphi^*(x))(\nu_1(x) - r_0\nu_0(x)) = 0(\mathrm{mod}\mu).$$

故 $\varphi^*(x)=\widetilde{\varphi}(x)(\mathrm{mod}\mu)$于$\{\nu_1(x)\neq r_0\nu_0(x)\}$上,即 φ^* 满足要求 (b).

令 $A=\{x\in E;\nu_1(x)=r_0\nu_0(x)\}$,则

$$E_0\varphi^*(\xi) = E_\mu(\widetilde{\varphi}\nu_0 \cdot 1_{A^c}) + E_\mu(\varphi^*\nu_0 1_A) \leqslant \alpha = E_0\widetilde{\varphi}(\xi);$$
$$E_1\varphi^*(\xi) = E_\mu(\widetilde{\varphi}\nu_1 1_{A^c}) + E_\mu(\varphi^*\nu_1 1_A) = E_1\widetilde{\varphi}(\xi).$$

于是

$$E_\mu(\varphi^*\nu_1 1_A)=E_\mu(\widetilde{\varphi}\cdot\nu_1 1_A),E_\mu(\widetilde{\varphi}\cdot\nu_0 1_A)\leqslant E_\mu(\widetilde{\varphi}\nu_0 1_A) \quad (6.2\text{-}1)$$
$$r_0E_\mu(\varphi^*\nu_0 1_A) = r_0E_\mu(\widetilde{\varphi}\cdot\nu_0\cdot 1_A). \quad (6.2\text{-}2)$$

如果 (6.2-1)中第二式的等号成立,则必有

$$E_0\varphi^*(\xi) = \alpha.$$

从而,φ^* 满足要求(a). 如果(6.2-1)中第二式严格不等,则(6.2-2)式只有在 $r_0=0$ 时才能成立. 这时

$$E_0\varphi^*(\xi) < \alpha,\ E_1\varphi^*(\xi) = E_1(\widetilde{\varphi}(\xi)1_{A^c}) = 1.$$

这正是结论(iii)中要排去的例外情形.

推论 6.2.1 设 $\alpha\in(0,1)$,β 为水平 α 下"检验 H 反对 K"的最大权. 若 $P_0\neq P_1$,则 $\alpha<\beta$.

证 显然,$\varphi(x)\equiv\alpha\in\Phi_\alpha$. 按假设,有

$$E_1\varphi(\xi) \leqslant \beta,\text{即 } \alpha \leqslant \beta.$$

若 $\alpha=\beta$,则 $\varphi(x)\equiv\alpha$ 为水平 α 下的最强检验函数,按定理 2.1 (iii),有

$$\alpha = \varphi(x) = \widetilde{\varphi}(x)(\mathrm{mod}\mu) \text{ 于} \{\nu_1(x) \neq r_0\nu_0(x)\} \text{ 上}.$$

这里 $\mu=\widetilde{P}_0+\widetilde{P}_1$,按 $\widetilde{\varphi}$之定义及 $0<\alpha<1$,有

$$\mu(\nu_1(x) \neq r_0\nu_0(x)) = 0$$
$$\Rightarrow \mu(\nu_1(x) = r_0\nu_0(x)) = 1$$

$\Rightarrow 1 = E_\mu \nu_1 = r_0 E_\mu \nu_0 = r_0$

$\Rightarrow \nu_1(x) = \nu_0(x) (\bmod \mu)$ 于 E 上

$\Rightarrow P_0 = P_1$. 这与假设矛盾. 故 $\alpha < \beta$.

§ 6.3　检验两假设问题

设 $\xi = (\xi_n, n \in \mathcal{N})$: ξ_n 是 $(\Omega, \mathcal{F}) \to (\mathcal{R}, \mathcal{B})$ 的随机元, $\mathcal{N}_1 = \{1, 2, \cdots\}$, $\mathcal{F}^\xi = \sigma(\xi)$. $\mathcal{P} = (P^i, i = 0, 1)$, 为 $(\Omega, \mathcal{F}_\xi)$ 上的概率测度族. "假设 H_i" $\triangleq$ "ξ 服从分布 P^i" $(i = 0, 1)$. 决策空间 $D = \{0, 1\}$. θ 为 $(\Omega, \mathcal{F}) \to \Theta = \{0, 1\}$ 的可测映射. π_0 为 $(\Omega, \mathcal{F}) \to ([0, 1], \mathcal{B} \cap [0, 1])$ 的随机元, 且具有概率分布 λ. 在 $(\Omega, \mathcal{F})$ 上引入概率测度 P, 满足下列要求:

(1) $P(\theta = 1/\pi_0) = \pi_0$. $P(\theta = 0/\pi_0) = 1 - \pi_0$;

(2) $P(\theta = i, A/\pi_0) = P(\theta = i/\pi_0)$. $P^i(A)$ $(A \in \mathcal{F}^\xi, i = 0, 1)$;

(3) $P|_{\sigma(\pi_0)} = \lambda$. 如果 λ 是集中在 $\pi \in [0, 1]$ 上的分布, 则记 P 为 P_π.

定义下列符号:

$(\mathcal{F}^\xi_\cdot) = (\mathcal{F}^\xi_n, n \in \mathrm{N})$: $\mathcal{F}^\xi_0 = \{\phi, \Omega\}$, $\mathcal{F}^\xi_n = \sigma(\xi_1, \xi_2, \cdots, \xi_n)$ $(n \geqslant 1)$.

$$\mathcal{M}^\xi = \{\tau : \tau \text{ 为有限的}(\mathrm{N}, \mathcal{F}^\xi_\cdot)\text{-时}\};$$

$$\mathcal{D}^\xi_\tau = \{d : d \text{ 取值于 } D, \text{且 } \mathcal{F}^\xi_\tau\text{-可测}\}; \ (\tau \in \mathcal{M}^\xi).$$

定义 6.3.1　设

$$\Delta = \{\delta : \delta = (\tau, d), \tau \in \mathcal{M}^\xi, d \in \mathcal{D}^\xi_\tau\}.$$

则称 Δ 为决策规则空间. $\delta \in \Delta$ 称为决策规则或决策函数, 其中 τ 表示停止观测的时刻. $\{d = i\}$ 表示接受假设 H_i. $(i = 0, 1)$.

考虑两种类型的损失. 第一类损失由观测代价引起. 假定观测一次的代价为 $c > 0$, 则由停止规则 $\tau \in \mathcal{M}^\xi$ 引起的损失为 $c\tau$. 第二类损失由决策 $d \in \mathcal{D}^\xi_\tau$ 引起, 偏差函数定义如下:

$$L(\theta, d) = a 1_{\{\theta = 1, d = 0\}} + b 1_{\{\theta = 0, d = 1\}},$$

其中a、b为正常数. 在决策规则$\delta=(\tau,d)\in\Delta$下所引起的风险(平均损失)为

$$\rho_\pi(\delta)=E_\pi(c\cdot\tau)+E_\pi L(\theta,d).$$

其中E_π为关于P_π的期望算子. 令

$$\alpha(\delta)=P'(d=0),\ \beta(\delta)=P^0(d=1),$$

则
$$\rho_\pi(\delta)=a\pi\alpha(\delta)+b(1-\pi)\beta(\delta)+cE_\pi\tau,$$

$$E_\pi\tau=\pi E'\tau+(1-\pi)E^0\tau.$$

其中E^i表示关于P^i的期望算子($i=0,1$).

定义 6.3.2 如果存在一个决策规则$\delta^*=(\tau^*,d^*)\in\Delta$,使得

$$\rho_\pi(\delta^*)=\inf\{\rho_\pi(\delta):\delta\in\Delta\}(\forall\pi\in[0,1]),$$

则称δ^*为关于Δ的π-Bayes 最优决策规则.

注 1 这里介绍的概念实则为问题Ⅴ的一种翻版. 不同点在于损失与风险均有明确的表示. 本章后面的几节将只讨论与π-Bayes 最优决策有关的问题.

§ 6.4 π-Bayes 最优决策问题的转化

本节讨论由π-Bayes 最优决策到最优停止的转化问题.

定义 6.4.1 设$\zeta=(\zeta_n,n\in\mathrm{N})$, $\eta=(\eta_n,n\in\mathrm{N})$为定义在概率空间$(\Omega,\mathscr{F},P)$上,且分别取值于$(X,\mathscr{X})$,$(Y,\mathscr{Y})$的随机序列. $(\mathscr{F}^\zeta_\cdot)=(\mathscr{F}^\zeta_n,n\in\mathscr{N})$,其中$\mathscr{F}^\zeta_n=\sigma(\zeta_k,0\leqslant k\leqslant n)$. 若

(i)η_n是$\mathscr{F}^\zeta_n$-可测的,即$\eta_n^{-1}(B)\in\mathscr{F}^\zeta_n(\forall\ B\in\mathscr{Y})$, $(\forall n\in\mathrm{N})$;

(ii)对$\forall\ n\geqslant1$,$\exists$ 映射$\varphi_n:(Y\times X,\mathscr{Y}\times\mathscr{X})\to(Y,\mathscr{Y})$,

使
$$\eta_n(\omega)=\varphi_n(\eta_{n-1}(\omega),\zeta_n(\omega))\ (\mathrm{mod}P).$$

则称η为一个关于$(\mathscr{F}^\zeta_\cdot)$的迁移统计系.

引理 6.4.1 设η为定义 6.4.1 中的迁移统计系. 若对$\forall\ n\in\mathrm{N}$,有

$$P(\zeta_{n+1}\in B|\mathscr{F}^\zeta_n)=P(\zeta_{n+1}\in B|\eta_n)\ (\forall\ B\in\mathscr{X}),\qquad(6.4\text{-}1)$$

则随机元η为关于P的一个 Markov 随机函数,即

$$P(\eta_{n+1}\in A|\mathscr{F}_n^{\zeta})=P(\eta_{n+1}\in A|\eta_n)\ (\forall\ A\in\mathscr{Y})\ (\forall\ n\in \mathrm{N}).$$
(6.4-2)

证 (6.4-2)式等价于:对任意有界可测映射

$$f:(Y,\mathscr{Y})\to(\mathrm{R},\mathscr{B}),$$

有 $E(f(\eta_{n+1})|\mathscr{F}_n^{\zeta})=E(f(\eta_{n+1})|\eta_n)\ (\mathrm{mod}P)\ (\forall\ n\in\mathrm{N})$

(6.4-3)

设 $\varphi(y,x)=\sum_{k=1}^{m}g_k(y)1_{B_k}(x)\ (B_k\in\mathscr{X})$,则

$$\begin{aligned}E(\varphi(\eta_n,\zeta_{n+1})|\mathscr{F}_n^{\zeta})&=\sum_{k=1}^{m}E(g_k(\eta_n)\cdot 1_{B_K}(\zeta_{n+1})|\mathscr{F}_n^{\zeta})\\&=\sum_{k=1}^{m}g_k(\eta_n)E(1_{B_k}(\zeta_{n+1})|\mathscr{F}_n^{\zeta})\\&=\sum_{k=1}^{m}g_k(\eta_n)E(1_{B_k}(\zeta_{n+1})|\eta_n)\\&=E(\varphi(\eta_n,\zeta_{n+1})|\eta_n).\end{aligned}$$

因此,(6.4-3)式对 f 为简单函数时成立.对一般情形,可通过简单函数逼近,并利用 Lebesque 控制收敛定理即可得所要求的结论.

现在,设 $\xi=(\xi_n,n\in\mathrm{N}_1)$:$\xi_n$ 是$(\Omega,\mathscr{F})\to(\mathrm{R},\mathscr{B})$的随机元.$P^i$ $(i=0、1)$和 P 均按§3 中所设.假设 ξ 关于 P^i 独立同分布.易证,ξ 在$(\Omega,\mathscr{F},P)$上也是独立同分布的.假定密度 $\nu_i(x)=\dfrac{\mathrm{d}P^i(\xi_1\leqslant x)}{\mathrm{d}x}$ $(x\in\mathscr{R})$存在.

令 $\mathscr{F}_0=\sigma(\pi_0)$, $\mathscr{F}_n=\sigma(\pi_0;\xi_1,\xi_2,\cdots,\xi_n)$.

$$\pi_0=P(\theta=1|\mathscr{F}_0),\ \pi_n=P(\theta=1|\mathscr{F}_n).\qquad(6.4\text{-}4)$$

按可测函数的表现定理,∃ 多元 Borel 函数 φ,使得

$$\pi_n=\varphi(\pi_0,\xi_1,\xi_2,\cdots,\xi_n).\qquad(6.4\text{-}5)$$

记 $\pi_n^{\pi}=\pi_n|_{\pi_0=\pi}$,则 π_n^{π} 是 $\mathscr{F}_n^{\xi}$-可测的$(n\in\mathrm{N})$.

引理 6.4.2 $\Pi=(\pi_n,n\in\mathrm{N})$有如下表现:

$$\pi_{n+1} = \frac{\pi_0 \prod_{k=1}^{n+1} \nu_1(\xi_k)}{\pi_0 \prod_{k=1}^{n+1} \nu_1(\xi_k) + (1-\pi_0) \prod_{k=1}^{n+1} \nu_0(\xi_k)}$$

$$= \frac{\pi_n \nu_1(\xi_{n+1})}{\pi_n \nu_1(\xi_{n+1}) + (1-\pi_n)\nu_0(\xi_{n+1})}. \qquad (6.4\text{-}6)$$

证 先证

$$P(\theta=1, \xi_{n+1} \in B | \mathscr{F}_n) = \pi_n P'(\xi_{n+1} \in B) (\mathrm{mod} P). \qquad (6.4\text{-}7)$$

事实上,令 $\mathscr{A}_n = \{A_0 \cap A' : A_0 \in \sigma(\pi_0), A' \in \mathscr{F}_n^\xi\}$,则 $\mathscr{A}_n$ 为 Ω 上的一个 π-类,$\Omega \in \mathscr{A}_n$,且 $\mathscr{F}_n = \sigma(\mathscr{A}_n)$. 于是,(6.4-7)式等价于:对$\forall\ A \in \mathscr{A}_n$,有

$$\int_A P(\theta=1, \xi_{n+1} \in B | \mathscr{F}_n) \mathrm{d}P = \int_A \pi_n \cdot P'(\xi_{n+1} \in B) \mathrm{d}P. \qquad (6.4\text{-}8)$$

注意到 ξ 的独立性及 P 之定义,有

$$\int_A P(\theta=1, \xi_{n+1} \in B | \mathscr{F}_n) \mathrm{d}P = \int_{A_0 \cdot A'} 1_{\{\theta=1, \xi_{n+1} \in B\}} \mathrm{d}P$$

$$= E(1_{A_0} P(\theta=1, \xi_{n+1} \in B, A' | \pi_0)).$$

$$= E(1_{A_0} \cdot \pi_0 \cdot P'(\xi_{n+1} \in B, A'))$$

$$= P'(\xi_{n+1} \in B) E(1_{A_0} \pi_0 P'(A'))$$

$$= P'(\xi_{n+1} \in B) E(1_{A_0} P(\theta=1, A' | \pi_0))$$

$$= P'(\xi_{n+1} \in B) \int_{A_0 \cap A'} 1_{\{\theta=1\}} \mathrm{d}P$$

$$= \int_A P'(\xi_{n+1} \in B) \cdot \pi_n \mathrm{d}P,$$

故(6.4-8)式成立.

同理可证

$$P(\theta=0, \xi_{n+1} \in B | \mathscr{F}_n) = (1-\pi_n) P^0(\xi_{n+1} \in B) \ (\mathrm{mod} P). \qquad (6.4\text{-}9)$$

将(6.4-8)和(6.4-9)式结合起来,得

$$P(\xi_{n+1} \in B|\mathscr{F}_n) = \pi_n P^1(\xi_{n+1} \in B) + (1 - \pi_n)P^0(\xi_{n+1} \in B). \tag{6.4-10}$$

另一方面，有

$$\begin{aligned} P(\theta = 1, \xi_{n+1} \in B|\mathscr{F}_n) &= E(P(\theta = 1, \xi_{n+1} \in B|\mathscr{F}_{n+1})|\mathscr{F}_n) \\ &= E(1_{\{\xi_{n+1}\in B\}}P(\theta = 1|\mathscr{F}_{n+1})|\mathscr{F}_n) \\ &= E(1_{\{\xi_{n+1}\in B\}}\pi_{n+1}|\mathscr{F}_n). \end{aligned} \tag{6.4-11}$$

联合(6.4-8)，(6.4-10)及(6.4-11)诸式，可得

$$\frac{E(1_{\{\xi_{n+1}\in B\}}\pi_{n+1}|\mathscr{F}_n)}{P(\xi_{n+1}\in B|\mathscr{F}_n)} = \frac{\pi_n P^1(\xi_{n+1}\in B)}{\pi_n P^1(\xi_{n+1}\in B) + (1-\pi_n)P^0(\xi_{n+1}\in B)}. \tag{6.4-12}$$

现在，让 $x\in B$，且 $B\downarrow\{x\}$，在(6.4-12)式两边取极限，然后，将 x 改成 ξ_{n+1}，即得(6.4-6)中的第二等式. 利用此式反复迭代，即得(6.4-6)中的第一等式.

引理 6.4.3 $\Pi=(\pi_n,\mathscr{F}_n,n\in\mathbb{N})$ 为一个迁移统计系，且为齐次 Markov 随机函数.

证 按(6.4-6)式，立即可验证定义 6.4.1 中的条件(i)和(ii)，于是，Π 为一个迁移统计系.

按引理 6.4.1，Π 的 Markov 性由下式保证：

$$P(\xi_{n+1} \in B|\mathscr{F}_n) = P(\xi_{n+1} \in B|\pi_n) \quad (\forall\, B \in \mathscr{B},\ n \in \mathbb{N}).$$

此式由(6.4-10)式推出.

最后证明：Π 的齐次性. 设 f 为任意有界 Borel 函数，则

$$\begin{aligned} E(f(\pi_{n+1})|\pi_n)|_{\pi_n=\pi} &= E(f(\varphi(\pi_n,\xi_{n+1}))|\pi_n)|_{\pi_n=\pi} \\ &= E_\pi(f(\varphi(\pi,\xi_{n+1}))) = E_\pi(f(\varphi(\pi,\xi_1))) \\ &= E_\pi f(\pi_1), \end{aligned}$$

故 $$E(f(\pi_{n+1})|\pi_n) = E_{\pi_n}f(\pi_1) \ (\mathrm{mod}P).$$

引理 6.4.4 对 $\forall\ \pi\in[0,1]$，

(i) $\Pi=(\pi_n,\mathscr{F}_n,n\in\mathbb{N})$ 关于 P_π 为齐次 Markov 链；

(ii) $\Pi_\pi=(\pi_n^\pi,\mathscr{F}_n^\xi,n\in\mathbb{N})$ 关于 P_π 为 Doob 鞅.

证　由引理 6.4.3 及 π_τ^π 之定义得出.

定理 6.4.1　设 $\tau\in\mathscr{M}^\xi, d^*=\begin{cases}1, & a\pi_\tau^\pi\geqslant b(1-\pi_\tau^\pi),\\ 0, & \text{否则}.\end{cases}$ 则

(i) $\delta^*=(\tau,d^*)\in\Delta$;

(ii) $\rho_\pi(\delta^*)\leqslant\rho_\pi(\delta)$，其中 $\delta=(\tau,d), d\in\mathscr{D}_\tau^\xi$.

$$\rho_\pi(\delta)=\rho_\pi(\tau,d)=E_\pi(c\cdot\tau+L(\theta,d))$$
$$=cE_\pi\tau+a\pi\alpha(\delta)+b(1-\pi)\beta(\delta);$$
$$\alpha(\delta)=P'(d=0),\ \beta(\delta)=P^0(d=1).$$

(iii) $\rho_\pi=\inf\{E_\pi(c\cdot\tau+L(\theta,d)):\delta=(\tau,d)\in\Delta\}$

$$=\inf\{E_\pi(c\cdot\tau+g(\pi_\tau^\pi)):\tau\in\mathscr{M}^\xi\},$$

其中 $g(\pi)=(a\pi)\wedge(b(1-\pi))\ (\pi\in[0,1])$.

证　显然，π_τ^π 是 $\mathscr{F}_\tau^\xi$-可测的. 因此，$d^*\in\mathscr{D}_\tau^\xi$. 从而 $\delta^*\in\Delta$. 此即结论(i).

结论(ii)、(iii)分以下几个步骤证明. 先证

$$\pi_\infty^\pi=P_\pi(\theta=1|\mathscr{F}^\xi)=\pi\frac{\mathrm{d}P'}{\mathrm{d}P_\pi}\ (\mathrm{mod}P_\pi)\ (\forall\ \pi\in[0,1]).$$

事实上，按引理 6.4.4，π_∞^π 存在，且 $\pi_\infty^\pi=P(\theta=1|\mathscr{F}^\xi)$. 对 $\forall\ A\in\mathscr{F}^\xi$，有

$$\int_A\pi_\infty^\pi\mathrm{d}P_\pi=P_\pi(\theta=1,A)=\pi\cdot P'(A)=\int_A\pi\frac{\mathrm{d}P'}{\mathrm{d}P_\pi}\mathrm{d}P_\pi$$

$$\Rightarrow\pi_\infty^\pi=\pi\frac{\mathrm{d}P'}{\mathrm{d}P_\pi}\ (\mathrm{mod}P_\pi).$$

次证　$$P_\pi(\theta=1,d=0)=E_\pi((1-d)\pi_\tau^\pi),$$

$$P_\pi(\theta=0,d=1)=E_\pi(d\cdot(1-\pi_\tau^\pi)).$$

事实上

$$P_\pi(\theta=1,d=0)=\pi P'(d=0)=\pi(1-P'(d=1))$$
$$=\pi E'(1-d)=\pi E_\pi\left((1-d)\frac{\mathrm{d}P'}{\mathrm{d}P_\pi}\right)$$
$$=\pi E_\pi\left((1-d)E_\pi\left(\frac{\mathrm{d}P'}{\mathrm{d}P_\pi}\Big/\mathscr{F}_\tau^\xi\right)\right)$$

$$= E_\pi((1-d)\pi_\tau^\pi).$$

同理可得 $P_\pi(\theta=0,d=1)=E_\pi(d\cdot(1-\pi_\tau^\pi))$.

再次证(ii). 设 $\delta=(\tau,d)\in\Delta$. 则按上述结果,有

$$\begin{aligned}\rho_\pi(\delta) &= E_\pi(c\cdot\tau+L(\theta,d)) \\ &= cE_\pi\tau + aP_\pi(\theta=1,d=0)+bP_\pi(\theta=0,d=1) \\ &= cE_\pi\tau + E_\pi(a\pi_\tau^\pi(1-d)+b(1-\pi_\tau^\pi)d) \\ &\geqslant E_\pi(c\cdot\tau+g(\pi_\tau^\pi)) = \rho_\pi(\delta^*).\end{aligned}$$

最后证(iii). 按结论(ii),有

$$\begin{aligned}\rho_\pi &= \inf\{\rho_\pi(\delta^*):\tau\in\mathscr{M}^\xi\} \\ &= \inf\{\rho_\pi(\delta):\delta\in\Delta\}.\end{aligned}$$

定理 6.4.2 设 $\mathscr{M}(\mathscr{F}.)=\{\tau:\tau$ 为有限$(\mathscr{N},\mathscr{F}.)$-时$\}$,

$$\mathscr{M}'(\mathscr{F}.)=\{\tau\in\mathscr{M}(\mathscr{F}.):E_\pi\tau<\infty.\ (\forall\ \pi\in[0,1])\}.$$

$$g(\pi)=(a\pi)\wedge(b\cdot(1-\pi))\ (\pi\in[0,1]).$$

则定义 6.3.2 中的 π-Bayes 最优决策问题可转化成关于齐次 Markov 链 $\Pi=(\pi_n,\mathscr{F}_n,n\in\mathrm{N})$的最优停时问题,即

$$\rho(\pi)=\inf\{E_\pi(c\cdot\tau+g(\pi_\tau)):\tau\in\mathscr{M}'(\mathscr{F}.)\}\ (\forall\ \pi\in[0,1]) \tag{6.4-13}$$

证 $\tau\in\mathscr{M}(\mathscr{F}.)$表明:$\tau$ 是$\mathscr{F}_\infty$-可测的. 因此,有

$$\tau=\tau(\pi_0;\xi_1,\xi_2,\cdots).$$

令 $\tau'=\tau|_{\pi_0=\pi}$,则 $\tau'\in\mathscr{M}^\xi$,且 $E_\pi\tau=E_\pi\tau'$,于是

$$E_\pi g(\pi_\tau)=E_\pi g(\pi_{\tau'}^\pi).$$

反之,对$\forall\ \tau'\in\mathscr{M}^\xi$,让 $\tau=\tau'$,则 $\tau\in\mathscr{M}(\mathscr{F}.)$.

上述分析表明:

$$\begin{aligned}&\inf\{E_\pi(c\cdot\tau+g(\pi_\tau)):\tau\in\mathscr{M}'(\mathscr{F}.)\} \\ &\qquad= \inf\{E_\pi(c\cdot\tau+g(\pi_\tau^\pi)):\tau\in\mathscr{M}^\xi\}.\end{aligned}$$

按定理 6.4.1, $\rho(\pi)=\rho_\pi$.

推论 6.4.1 设 $\tau^*\in\mathscr{M}'(\mathscr{F}.)$,且

$$\rho(\pi)=E_\pi(c\tau^*+g(\pi_{\tau^*}))\ (\forall\ \pi\in[0,1]),$$

则在定义 6.3.2 中的 π-Bayes 最优决策问题之解 δ^* 为

$$\delta^* = (\tau^*, d^*),\ d^* = 1_{\{\pi_{\tau^*} \geqslant \frac{b}{a+b}\}}$$

且决策规则 δ^* 所引起的两类偏差为

$$\alpha_\pi(\delta^*) = P_\pi(d^* = 0/\theta = 1) = \frac{P_\pi(\theta = 1, d^* = 0)}{P_\pi(\theta = 1)}$$

$$= \frac{1}{\pi} E_\pi((1 - d^*)\pi_{\tau^*});$$

$$\beta_\pi(\delta^*) = P_\pi(d^* = 1/\theta = 0) = \frac{P_\pi(\theta = 0, d^* = 1)}{P_\pi(\theta = 0)}$$

$$= \frac{1}{1-\pi} E_\pi(d^*(1 - \pi_{\tau^*})).$$

注 1 将(6.4-13)式两边改号即可得 π-Bayes 最优决策问题的另一种形式：

$$\bar{\rho}(\pi) = -\rho(\pi) = \sup\{E_\pi(\bar{g}(\pi_\tau) - c\tau) : \tau \in \mathscr{M}'(\mathscr{F}.)\},$$

其中 $\bar{g}(\pi) = -g(\pi)$. 因此，π-Bayes 最优决策问题被转化为一个具有固定观测代价 c，折扣率为 1 的 Markov 最优停时问题. 该问题的一般结果将在下一节介绍.

§6.5 有观测代价的最优停止

设 $X = (X_n, \mathscr{F}_n, n \in \mathrm{N})$ 是 $(\Omega, \mathscr{F}) \to (E, \mathscr{E})$ 的随机元序列. 其中 $(\mathscr{F}.)$ 为 $\mathscr{F}$ 中的 σ-代数流，X_n 是 $\mathscr{F}_n$-可测的. 对 $\forall\ x \in E$，X 关于 P_x 为齐次 Markov 链，且初始状态为 x. 定义下列符号：

$$g = g(x)\ (x \in E) \text{——目标函数};$$

$$c = c(x)\ (x \in E) \text{—— 观测代价函数};$$

$$\psi_{0,0}(g) = g(x_0);$$

$$\psi_{n,K}(g) = \alpha^k \cdot g(X_{n+k}) - \sum_{s=n}^{n+k-1} \alpha^{s-n} c(X_s),\ (n \geqslant 1, k \geqslant 1); \tag{6.5-1}$$

$\alpha \in (0,1]$——折算率；

$$\mathscr{M} = \{\tau : \tau \text{ 为有限}(\mathscr{N}, \mathscr{F}_0)\text{-时}\};$$

$$\mathscr{M}(\alpha,c) = \{\tau \in \mathscr{M} : E_x\Big(\sum_{k=0}^{\tau-1} \alpha^k c(X_k) 1_{(\tau \geq 1)}\Big) < \infty \ (\forall\, x \in E)\}. \tag{6.5-2}$$

$$S(x) = \sup\{E_x \psi_{0,\tau}(g) : \tau \in \mathscr{M}_{(\alpha,c)}\} \ (\text{报酬函数}). \tag{6.5-3}$$

本节讨论$(\psi_{0,n}(g), \mathscr{F}_n, n \in \mathrm{N})$的最优停止问题. 这里$\psi_{0,n}(g)$可以解释为从时刻0开始至时刻$n$的实际输出结果.

定义 6.5.1 设$\varepsilon \geqslant 0$,若$\tau_\varepsilon \in \mathscr{M}(\alpha,c)$满足要求:

$$S(x) - \varepsilon \leqslant E_x \psi_{0,\tau_\varepsilon}(g) \ (\forall\, x \in E), \tag{6.5-4}$$

则称τ_ε为g的(ε,S)-最优停时,简称τ_0为g的最优停时.

定义 6.5.2 设$f \in B$(定义见第五章). 若

$$f(x) \geqslant g(x) \vee (\alpha T f(x) - c(x)) \ (x \in E), \tag{6.5-5}$$

其中算子T定义如下:

$$Tf(x) = E_x f(X_1).$$

则称f为g的(α,c)-过剩主部.

下面讨论(α,c)-过剩主部的某些基本性质. 假设$0 < \alpha \leqslant 1$. $g \in \mathscr{B}$,且有界K. $c(x) \geqslant 0$,且$E_x c(X_n) < \infty$ $(\forall\, n \in \mathrm{N})$. 定义算子$Q_{(\alpha,c)}$:

$$Q_{(\alpha,c)} f(x) = f(x) \vee (\alpha T f(x) - c(x)). \tag{6.5-6}$$

引理 6.5.1

$$Q^n_{(\alpha,c)} g(x) \leqslant Q^{n+1}_{(\alpha,c)} g(x); \tag{6.5-7}$$

$$Q^n_{(\alpha,c)} g(x) = g(x) \vee (\alpha T Q^{n-1}_{(\alpha,c)} g(x) - c(x)). \tag{6.5-8}$$

证 由算子$Q_{(\alpha,c)}$之定义立即得(6.5-7)式.

按(6.5-7)式及算子T的保序性,有

$$\alpha T Q^{n-1}_{(\alpha,c)} g(x) - c(x) \leqslant \alpha T Q^n_{(\alpha,c)} g(x) - c(x). \tag{6.5-9}$$

反复应用(6.5-7)式及(6.5-9)式,可推出(6.5-8)式.

引理 6.5.2 令

$$v(x) = \lim_{n\to\infty} Q^n_{(\alpha,c)} g(x), \tag{6.5-10}$$

则(i)v为g的最小(α,c)-过剩主部,且

$$|v(x)| \leqslant \sup\{|g(x)|; x \in E\} = K. < \infty.$$

(ii)v 满足如下递归方程：

$$v(x) = g(x) \vee (\alpha Tv(x) - c(x)). \qquad (6.5\text{-}11)$$

证 若 $f_1(x) \leqslant f_2(x)$ $(x \in E)$，则由 $Q_{(\alpha,c)}$ 之定义，易知

$$Q^n_{(\alpha,c)} f_1(x) \leqslant Q^n_{(\alpha,c)} f_2(x).$$

$$\Rightarrow Q^n_{(\alpha,c)} |g(x)| \leqslant Q^n_{(\alpha,c)} K = K$$

$$\Rightarrow |v(x)| \leqslant K.$$

按引理 6.5.1，v 存在. 应用 Lebesque 有界收敛定理，可推出，v 满足方程(6.5-11).

为证结论(i)，设 f 为 g 的任意(α, c)-过剩主部，则按定义 6.5.2，对$\forall$ $n \in \mathrm{N}$，有

$$Q^n_{(\alpha,c)} f(x) \geqslant Q^n_{(\alpha,c)} g(x).$$

由(6.5-5)式，可得

$$f(x) \geqslant f(x) \vee (\alpha Tf(x) - c(x)) = Q_{(\alpha,c)} f(x)$$

$$\Rightarrow f(x) \geqslant Q^n_{(\alpha,c)} f(x)$$

$$\Rightarrow f(x) \geqslant Q^n_{(\alpha,c)} g(x)$$

$$\Rightarrow f(x) \geqslant v(x). \ (\forall\ x \in E).$$

引理 6.5.3 对$\forall$ $x \in E$，$(v(X_n), \mathscr{F}_n, n \in \mathrm{N})$关于 P_x 为(α, c)-上鞅，即

$$E_x(\alpha v(X_{n+1})/\mathscr{F}_n) - c(X_n) \leqslant v(X_n) \ (\mathrm{mod} P_x) \ (\forall\ n \in \mathrm{N}).$$

$(\psi_{0,n}(v), \mathscr{F}_n, n \in N)$关于 P_x 为上鞅，即

$$E_x(\psi_{0,n+1}(v)/\mathscr{F}_n) \leqslant \psi_{0,n}(v) \ (\mathrm{mod} P_x) \ (\forall\ n \in \mathrm{N}).$$

此外 $$E_x(\psi_{0,\sigma}(v)/\mathscr{F}_\tau) \leqslant \psi_{0,\tau}(v) \ (\mathrm{mod} P_x), \qquad (6.5\text{-}12)$$

其中 σ、$\tau \in \mathscr{M}_{(\alpha,c)}$，且 $\tau \leqslant \sigma$.

证 按引理 6.5.2(i)，v 有界. 按引理 6.5.2(ii)及 X 的齐次 Markov 性，有

$$E_x(\alpha v(X_{n+1})/\mathscr{F}_n) - c(X_n) = E_{X_n}(\alpha v(X_1)) - c(X_n)$$

$$\leqslant v(X_n) \ (\mathrm{mod} P_x).$$

这表明，$(v(X_n),\mathscr{F}_n,n\in\mathscr{N})$为$(\alpha,c)$-上鞅. 由此，得

$$E_x(\psi_{n,k+1}(v)/\mathscr{F}_{n+k}) = \alpha^k E_x(\alpha\cdot v(X_{n+k+1})/\mathscr{F}_{n+k})$$

$$-\sum_{s=n}^{n+k-1}\alpha^{s-n}c(X_s) - \alpha^k c(X_{n+k})$$

$$= \alpha^k(\alpha Tv(X_{n+k}) - c(X_{n+k})) - \sum_{s=n}^{n+k-1}\alpha^{s-n}c(X_s)$$

$$\leqslant \alpha^k v(X_{n+k}) - \sum_{s=n}^{n+k-1}\alpha^{s-n}c(X_s)$$

$$= \psi_{n,k}(v)(\mathrm{mod}P_x).$$

这表明：$(\psi_{n,k}(v),\mathscr{F}_{n+k},k\in\mathrm{N})$关于$P_x$为上鞅($\forall\ n\in\mathrm{N}$).

显然，$\psi_{n,\cdot}(v)$为上控广义上鞅. 按定理4.5.5，得(6.5-12)式.

注 6.5.1 “$(v(X_n),\mathscr{F}_n,n\in\mathrm{N})$为关于$P_x$的$(\alpha,c)$-上鞅”等价于“对$\forall\ n\in\mathscr{N}$，$(\psi_{n,k}(v),\mathscr{F}_{n+k},k\in\mathrm{N})$为关于$P_x$的广义上鞅”.

由上述引理可得到一个很有用的不等式：

$$v(x)\geqslant E_x\psi_{0,\tau}(v)\ (\forall\ \tau\in\mathscr{M}_{(\alpha,c)}),\ x\in E).\quad(6.5\text{-}13)$$

引理 6.5.4 $\overline{\lim\limits_{n\to\infty}}\alpha^n v(X_n)=\overline{\lim\limits_{n\to\infty}}\alpha^n g(X_n)\ (\mathrm{mod}P_x)\ (x\in E)$.

证 由v为g的(α,c)-过剩主部可知

$$\overline{\lim}\alpha^n v(x_n)\geqslant\overline{\lim}\alpha^n g(x_n).$$

下证反向不等式. 为此，令

$$\varphi(x) = E_x(\sup\nolimits_{k\in N}\psi_{0,k}(g)).$$

则由X的齐次 Markov 性，得

$$\varphi(X_n) = E_{X_n}(\sup\nolimits_{k\in N}\psi_{0,k}(g))$$

$$= E_x(\sup\nolimits_{k\in N}\psi_{n,k}(g)/\mathscr{F}_n)\ (\mathrm{mod}P_x).$$

如果φ为g的(α,c)-过剩主部，即

$$\varphi(x)\geqslant g(x)\vee(\alpha T\varphi(x)-c(x))\ (\forall\ x\in E),\quad(6.5\text{-}14)$$

则由v的最小性可得

$$\alpha^n v(X_n)\leqslant\alpha^n\varphi(X_n) = E_x(\alpha^n\sup\nolimits_{k\in N}\psi_{n,k}(g)/\mathscr{F}_n)$$

$$\leqslant E_x(\sup\nolimits_{k\geqslant n}(\alpha^k\cdot g(X_k))/\mathscr{F}_n)\ (\mathrm{mod}P_x).$$

按推论 4.4.4.，得

$$\overline{\lim_{n\to\infty}}\alpha^n v(X_n)\leqslant \overline{\lim_{n\to\infty}} E_x(\sup_{k\geqslant n}(\alpha^k\cdot g(X_k))/\mathscr{F}_n)$$

$$=\overline{\lim_{k\to\infty}}\alpha^k g(X_k)\ (\mathrm{mod} P_x).$$

这就证实反向不等式成立.

剩下的问题就是证明(6.5-14).

$$\begin{aligned}\alpha T\varphi(x)&=\alpha E_x\varphi(X_1)=\alpha E_x E_x(\sup_{k\in N}\psi_{1,k}(g)/\mathscr{F}_1)\\&=\alpha E_x(\sup_{k\in N}\psi_{1,k}(g))\\&=\alpha E_x(\sup_{k\in N}(\alpha^k g(X_{1+k})-\sum_{s=1}^{k}\alpha^{s-1}c(X_s)))\\&\leqslant E_x(\sup_{k\in N}\psi_{0,k}(g))+E_x c(X_0)\\&=\varphi(x)+c(x).\end{aligned}$$

于是(6.5-14)式成立.

引理 6.5.5 令 $\tau_\varepsilon=\inf\{n\in N:\alpha^n v(X_n)\leqslant\alpha^n g(X_n)+\varepsilon\}$，则

(i) $\tau_\varepsilon\in\mathscr{M}_{(\alpha,c)}$，$P_X(\tau_\varepsilon<\infty)=1\ (\forall\ \varepsilon>0,x\in E)$；

(ii) $v(x)=E_x\psi_{0,n\wedge\tau_\varepsilon}(v)\ (\forall\ n\in \mathrm{N},x\in E,\varepsilon\geqslant 0)$；

(iii) $v(x)=E_x\psi_{0,\tau_\varepsilon}(v)\ (\forall\ x\in E,\varepsilon>0)$.

证 设 $\varepsilon>0$，则由 $0<\alpha\leqslant 1$，g 有界及引理 6.5.4 可知.

$$P_x(\tau_\varepsilon=\infty)\leqslant P_x(\overline{\lim}\alpha^n v(X_n)>\overline{\lim}\alpha^n g(X_n))=0$$

$$\Rightarrow P_x(\tau_\varepsilon<\infty)=1.$$

显然，对 $\forall\ \varepsilon\geqslant 0$，$n\wedge\tau_\varepsilon\in\mathscr{M}_{(\alpha,c)}(\forall\ n\in\mathrm{N})$，且

$$v(X_{n-1})>g(X_{n-1})\ 于(\tau_\varepsilon\geqslant n)\ 上.$$

按引理 6.5.2(ii)，有

$$v(X_{n-1})=\alpha Tv(x_{n-1})-c(X_{n-1})\ 于(\tau_\varepsilon\geqslant n)\ 上.$$

利用 $(\tau_\varepsilon\geqslant n)\in\mathscr{F}_{n-1}$ 及 X 的齐次 Markov 性，可推出：

$$E_x(v(X_n)1_{(\tau_\varepsilon\geqslant n)})=E_x(E_x(v(X_n)/\mathscr{F}_{n-1})1_{(\tau_\varepsilon\geqslant n)})$$

$$=E_x(Tv(X_{n-1})1_{(\tau_\varepsilon\geqslant n)})$$

$$\Rightarrow E_x(\psi_{0,n}(v)1_{(\tau_\varepsilon\geqslant n)})=E_x([(\alpha^n v(X_n)-\alpha^{n-1}c(X_{n-1}))$$

$$- \sum_{s=0}^{n-2}\alpha^s c(X_s)\Big]1_{(\tau_\varepsilon \geqslant n)}\Big)$$

$$= E_x\Big[(\alpha^{n-1}v(X_{n-1}) - \sum_{s=0}^{n-2}\alpha^s c(X_s))1_{(\tau_\varepsilon \geqslant n)}\Big]$$

$$= E_x(\psi_{0,n-1}(v) \cdot 1_{(\tau_\varepsilon \geqslant n)})$$

$$\Rightarrow E_x\psi_{0,n\wedge\tau_\varepsilon}(v) = E_x(\psi_{0,\tau_\varepsilon}(v)1_{(\tau_\varepsilon \leqslant n-1)}) + E_x(\psi_{0,n}(v)1_{(\tau_\varepsilon \geqslant n)})$$

$$= E_x(\psi_{0,\tau_\varepsilon}(v)1_{(\tau_\varepsilon \leqslant n-1)}) + E_x(\psi_{0,n-1}(v)1_{(\tau_\varepsilon \geqslant n)})$$

$$= E_x(\psi_{0,(n-1)\wedge\tau_\varepsilon}(v)) = \cdots = v(x). \Rightarrow \text{(ii)}.$$

按引理 6.5.2(i),v 以 K 为界. 按结论(ii),有

$$E_x\Big(\sum_{s=0}^{n\wedge\tau_\varepsilon - 1}\alpha^s c(X_s)\Big) = E_x(\alpha^{n\wedge\tau_\varepsilon}v(X_{n\wedge\tau_\varepsilon})) - v(x) \leqslant 2K < \infty.$$

注意,$P_x(\tau_\varepsilon<\infty)=1$ $(\varepsilon>0)$. 按 Fatou 引理,得

$$E_x\Big(\sum_{s=0}^{\tau_\varepsilon-1}\alpha^s. c(X_s)\Big) \leqslant 2k < \infty$$

$$\Rightarrow \tau_\varepsilon \in \mathscr{M}_{(\alpha,c)}.$$

故结论(i)成立.

对$\forall$ $\varepsilon>0$,按结论(i)、(ii),有

$$|\psi_{0,n\wedge\tau_\varepsilon}(v)| \leqslant K + \sum_{s=0}^{\tau_\varepsilon-1}\alpha^s. c(X_s) \quad (\forall\ n \in \mathrm{N}).$$

按 Lebesque 控制收敛定理,有

$$v(x) = \lim_{n\to\infty} E_x\psi_{0,n\wedge\tau_\varepsilon(v)}$$

$$= E_x(\lim_{n\to\infty}\psi_{0,n\wedge\tau_\varepsilon}(v)) = E_x\psi_{0,\tau_\varepsilon}(v).$$

此即结论(iii).

引理 6.5.6 设 τ_0 为引理 6.5.5 中 $\varepsilon=0$ 的情形.

(i)若 $P_x(\tau_0<\infty)=1$ $(\forall\ x\in E)$,则 $\tau_0\in\mathscr{M}_{(\alpha,c)}$,且

$$v(x) = E_x\psi_{0,\tau_0}(v). \quad (x\in E);$$

(ii)若 $P_x\Big(\sum_{s=0}^{\infty}\alpha^s \cdot c(X_s)=\infty\Big)=1$ $(\forall\ x\in E)$,则 $P_x(\tau_0<\infty)$

$=1\ (\forall\ x\in E)$.

证　当 $P_x(\tau_0<\infty)=1$ 时，在引理 6.5.5 中，让 $\varepsilon=0$，即得 (i).

现在采用反证法证明结论(ii)．假如有某个 $x_0\in E$，使得 $P_{x_0}(\tau_0<\infty)<1$，即 $P_{x_0}(\tau_0=\infty)>0$．则按引理 6.5.5，及 Fatou 引理，有

$$
\begin{aligned}
v(x_0)&=\overline{\lim_{n\to\infty}}E_{x_0}\psi_{0,n\wedge\tau_0}(v)\\
&\leqslant E_{x_0}\overline{\lim_{n\to\infty}}\psi_{0,n\wedge\tau_0}(v)=E_{x_0}\psi_{0,\tau_0}(v).
\end{aligned}
$$

按假设及 $\psi_{\cdot,\cdot}$ 之定义，必有 $E_{x_0}\psi_{0,\tau_0}(v)=-\infty$．故 $v(x_0)=-\infty$．这与 v 有界矛盾.

引理 6.5.7　$S(x)=v(x)\ (\forall\ x\in E)$.

证　按引理 6.5.2 和 6.5.3，对 $\forall\ \tau\in\mathscr{M}_{(\alpha,c)}$，有

$$
\begin{aligned}
v(x)\geqslant E_x\psi_{0,\tau}(v)&\geqslant E_x\psi_{0,\tau}(g)\ (\forall\ x\in E)\\
&\Rightarrow v(x)\geqslant S(x).
\end{aligned}
$$

另一方面，按引理 6.5.5，对 $\forall\ \varepsilon>0$，有

$$
\begin{aligned}
v(x)=E_x\psi_{0,\tau_\varepsilon}(v)&=E_x\Big(\alpha^{\tau_\varepsilon}\cdot v(X_{\tau_\varepsilon})-\sum_{s=0}^{\tau_\varepsilon-1}\alpha^s\cdot c(X_s)\Big)\\
&\leqslant E_x\Big(\alpha^{\tau_\varepsilon}g(X_{\tau_\varepsilon})-\sum_{s=0}^{\tau_\varepsilon-1}\alpha^s c(X_s)\Big)+\varepsilon\\
&=E_x\psi_{0,\tau_\varepsilon}(g)+\varepsilon\leqslant s(x)+\varepsilon\\
&\Rightarrow v(x)\leqslant S(x).
\end{aligned}
$$

定理6.5.1　设 $g,c\in\mathscr{B}$；$\|g\|=\sup\{|g(x)|:x\in E\}<\infty$；$c(x)\geqslant 0,E_xc(X_n)<\infty\ (\forall\ x\in E)$；$0<\alpha\leqslant 1$，则

(i) 报酬函数 s（由 6.5-3）式定义）为 g 的最小 (α,c) 过剩主部；

(ii) $s(x)=\lim_{n\to\infty}Q^n_{(\alpha,c)}g(x)\ (\forall\ x\in E)$，其中 $Q_{(\alpha,c)}$ 由(6.5-6)式定义；

(iii) s 满足如下方程：

$$\begin{cases} s(x) = g(x) \vee (\alpha T s(x) - c(x)); \\ \overline{\lim\limits_{n\to\infty}} \alpha^n S(X_n) = \overline{\lim\limits_{n\to\infty}} \alpha^n g(X_n) \ (\mathrm{mod} P_x) \end{cases} (x \in E),$$

(iv)令 $\tau_\varepsilon=\inf\{n\in N: \alpha^n S(X_n)\leqslant \alpha^n g(X_n)+\varepsilon\}$ $(\varepsilon\geqslant 0)$.

(1)若 $\varepsilon>0$,则 $\tau_\varepsilon\in\mathscr{M}_{(\alpha,c)}$,且为 g 的一个(ε,S)-最优停时.

(2)若 $P_x(\tau_0<\infty)=1$ $(\forall\ x\in E)$,则 $\tau_0\in\mathscr{M}_{(\alpha,c)}$,且为 g 的一个最优停时.

(3)若 $P_x\left(\sum\limits_{s=0}^{\infty}\alpha^s \cdot c(X_s)=\infty\right)=1$ $(\forall\ x\in E)$,则

$$P_x(\tau_0 < \infty) = 1 \ (\forall\ x \in E).$$

证 按引理 6.5.7,$v=S$. 从而,定理成立.

推论 6.5.1 设符号 $\Pi,\mathscr{F}.,g(\pi),\mathscr{M}'(\mathscr{F}.),\rho(\pi),c>0$ 均按定理 4.2 中定义,则

(i)$\bar{\rho}(\pi)=\lim Q^n_{(1,c)}\bar{g}(\pi)$ $(\pi\in[0,1])$,其中

$\bar{g}(\pi) = - g(\pi)$, $\bar{\rho}(\pi) = - \rho(\pi)$, $T\bar{g}(\pi) = E_\pi\bar{g}(\pi_1)$.

$Q_{(1,c)}\bar{g}(\pi) = \bar{g}(\pi) \vee (T\bar{g}(\pi) - c) = - g(\pi) \wedge (Tg(\pi) + c)$.

(ii) $\begin{cases} \bar{\rho}(\pi)=\bar{g}(\pi)\vee(T\bar{\rho}(\pi)-c); \\ \overline{\lim\limits_{n\to\infty}}\bar{\rho}(\pi_n)=\overline{\lim\limits_{n\to\infty}}\bar{g}(\pi_n) \ (\mathrm{mod}P_\pi) \end{cases}$ $(\pi\in[0,1])$.

(iii)$\tau_0\in\mathscr{M}'(\mathscr{F}.)$,且 $\bar{\rho}(\pi)=E_\pi(\bar{g}(\pi_{\tau_0})-c\cdot\tau_0)$ $(\pi\in[0,1])$.

其中 $\tau_0=\inf\{n\in\mathscr{N}:\bar{\rho}(\pi_n)=\bar{g}(\pi_n)\}$.

证 在定理 6.5.1 中,让 $\alpha=1,c(x)\equiv c>0,E=[0,1],g(\pi)=(a\pi)\wedge(b(1-\pi))$. 即可得此结论.

推论 6.5.2 在定义 6.3.2 中所表述的 π-Bayes 最优决策问题之解 $\delta^*=(\tau^*,d^*)$存在,且

$$\tau^* = \inf\{n \in \mathrm{N}: \rho(\pi_n) = g(\pi_n)\},$$

$$d^* = \begin{cases} 1, & \pi_{\tau^*} \geqslant \dfrac{b}{a+b}, \\ 0, & \pi_{\tau^*} < \dfrac{b}{a+b}. \end{cases}$$

证　由推论 6.5.1 立即得此结论.

引理 6.5.8　设 f、$g\in\mathscr{B}$，且有界；$c(x)\geqslant 0$，$E_x c(X_n)<\infty$ ($x\in E,n\in N$)，f 满足方程：

$$f(x)=g(x)\vee(\alpha Tf(x)-c(x))\quad(\forall\ x\in E).\tag{6.5-15}$$

$$\bar{\tau}_\varepsilon=\inf\{n\in \mathrm{N}:\alpha^n f(x_n)\leqslant\alpha^n g(X_n)+\varepsilon\}\quad(\varepsilon\geqslant 0).$$

则
$$f(x)=E_x\psi_{0,n\wedge\bar{\tau}_\varepsilon}(f)\quad(\forall\ x\in E,n\in \mathrm{N}).$$

证　本命题之证类似于引理 6.5.5(ii).

引理 6.5.9　若在引理 6.5.8 中补充假设

$$\overline{\lim_{n\to\infty}}\alpha^n f(X_n)=\overline{\lim_{n\to\infty}}\alpha^n g(X_n)\ (\mathrm{mod}P_x)\quad(x\in E).\tag{6.5-16}$$

则　$P_x(\bar{\tau}_\varepsilon<\infty)=1\ (\forall\ x\in E)$，且 $\bar{\tau}_\varepsilon\in\mathscr{M}_{(\alpha,c)}(\varepsilon>0)$.

证　类似于引理 6.5.5 之证.

定理 6.5.2　设 f、$g\in B$，且有界. $c(x)\geqslant 0$，$E_x c(x_n)<\infty$ ($\forall\ x\in E,n\in N$)，f 满足方程(6.5.15)，则 $f=s$ 的充要条件是：(6.5-16)式成立.

证　必要性即是定理 6.5.1(iii)，下证充分性.

按引理 6.5.8 和 6.5.9 及 Fatou 引理，有

$$\begin{aligned}f(x)&\leqslant\overline{\lim_{n\to\infty}}E_x\psi_{0,n\wedge\bar{\tau}_\varepsilon}(f)\leqslant E_x\psi_{0,\bar{\tau}_\varepsilon}(f)\\&\leqslant E_x\psi_{0,\bar{\tau}_\varepsilon}(g)+\varepsilon\leqslant s(x)+\varepsilon\ (\forall\ \varepsilon>0)\\&\Rightarrow f(x)=S(x).\end{aligned}$$

另一方面，f 满足方程(6.5-15)表明，f 为 g 的(α,c)-过剩主部. 从而，$S(x)\leqslant f(x)$.

联合起来得 $f=S$.

注 2　上述定理表明：定理 6.5.1(iii)中递归方程之解在 g 有界的条件下是唯一的.

推论 6.5.3　递归方程

$$\begin{cases} f(\pi) = g(\pi) \wedge (Tf(\pi) + c), \\ \varliminf_{n\to\infty} f(\pi_n) = \varliminf_{n\to\infty} g(\pi_n) \pmod{P_\pi} \end{cases} \quad (\pi \in [0,1]).$$

(其中 $c>0$，P_π 按§3中定义，$g(\pi)=(a\pi)\wedge(b(1-\pi))$)之解在有界 $\mathscr{B}\cap[0,1]$-可测函数范围内是唯一的.

证 这是定理 6.5.2 之特例.

§6.6 π-Bayes 最优决策规则的结构

本节将在推论 6.5.1～6.5.3 的基础上讨论 π-Bayes 最优决策的结构问题，在推论 6.5.1 中改变符号，有

$$Q_{(1,c)}g(\pi) = g(\pi) \wedge (Tg(\pi) + c). \tag{6.6-1}$$

$$\rho(\pi) = \lim_{n\to\infty} Q^n_{(1,c)} g(\pi) = g(\pi) \wedge (T\rho(\pi) + c); \tag{6.6-2}$$

$$\varliminf_{n\to\infty} \rho(\pi_n) = \varliminf_{n\to\infty} g(\pi_n) \pmod{P_\pi}; \tag{6.6-3}$$

$$g(\pi) = (a\pi) \wedge (b(1-\pi)). \tag{6.6-4}$$

本节采用的基本符号凡未加说明的均可在§6.3，§6.4中找到定义. 在引理 6.4.2 中已得到如下公式：

$$\pi_1 = \varphi(\pi_0, \xi_1) = \frac{\pi_0 \nu_1(\xi_1)}{\pi_0 \nu(\xi_1) + (1 - \pi_0)\nu_0(\xi_1)}. \tag{6.6-5}$$

引理 6.6.1

$$Tg(\pi) = E_\pi g(\pi_1) = \int_{\mathscr{R}} (a\pi\nu_1(x)) \wedge (b(1-\pi)\nu_0(x))\mathrm{d}x.$$

证

$$\begin{aligned} E_\pi g(\pi_1) &= \int_\Omega g(\pi_1) P_\pi(\mathrm{d}\omega) = \int_\Omega g(\varphi(\pi_0, \xi_1)) P_\pi(\mathrm{d}\omega) \\ &= \int_\Omega g(\varphi(\pi, \xi_1))(\pi P'(\mathrm{d}\omega) + (1-\pi)P^0(\mathrm{d}\omega)) \\ &= \int_{\mathscr{R}} g(\varphi(\pi, \xi_1))(\pi\nu_1(x) + (1-\pi)\nu^0(x))\mathrm{d}x; \end{aligned}$$

$$g(\pi_1) = \frac{a\pi_0\nu_1(\xi_1)}{\pi_0\nu_1(\xi_1) + (1-\pi_0)\nu_0(\xi_1)} \wedge \frac{b(1-\pi_0)\nu_0(\xi_1)}{\pi_0\nu_1(\xi_1) + (1-\pi_0)\nu_0(\xi_1)}$$

$$= \frac{(a\pi_0\nu_1(\xi_1)) \wedge (b(1-\pi_0)\nu_0(\xi_1))}{\pi_0\nu_1(\xi_1) + (1-\pi_0)\nu_0(\xi_1)}$$

$$\Rightarrow E_\pi g(\pi_1) = \int_{\mathscr{R}} (a\pi\nu_1(x)) \wedge (b(1-\pi)\nu_0(x)\mathrm{d}x.$$

引理 6.6.2 $g(\pi)$，$\rho(\pi)$，$T\rho(\pi)$均为上凸函数.

证 $g(\pi)$的上凸性是显然的（如图 6.6-1 所示）.

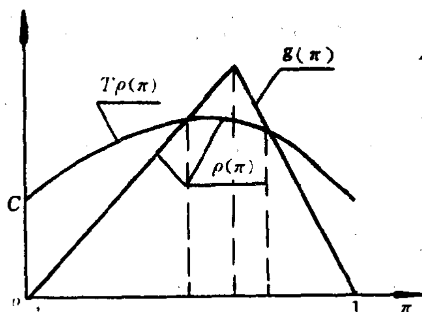


图 6.6-1 g、ρ、$T\rho$ 曲线

令 $g_x(\pi)=(a\pi\nu_1(x))\wedge(b(1-\pi)\nu_0(x))$，则对$\forall\ x\in\mathscr{R}$，$g_x(\pi)$为 $\pi\in[0,1]$的上凸函数. 于是，对$\forall\ \pi'$，$\pi''\in[0,1]$，及 λ_1、$\lambda_2\in[0,1]$，且 $\pi'<\pi''$，$\lambda_1+\lambda_2=1$，有

$$g_x(\lambda_1\pi' + \lambda_2\pi'') \geqslant \lambda_1 g_x(\pi') + \lambda_2 g_x(\pi'').$$

将此式两边关于 x 积分，得

$$Tg(\lambda_1\pi' + \lambda_2\pi'') \geqslant \lambda_1 Tg(\pi') + \lambda_2 Tg(\pi'').$$

故 $Tg(\pi)$为上凸函数. 从而，$Q_{(1,c)}g(\pi)$为上凸函数. 由归纳法立即可得 $Q^n_{(1,c)}g(\pi)$为上凸函数. 显然，上凸函数的极限为上凸函数，即 $\rho(\pi)$为上凸函数. 注意，对$\forall\ n\in\mathrm{N}$，$TQ^n_{(1,c)}g(\pi)$为上凸函数. 故其极限为上凸函数，即 $T\rho(\pi)$为上凸函数.

定理6.6.1 设ρ为递归方程（6.6-2）和（6.6-3）之解. 则 $\rho(\pi)$为 π-Bayes 最小风险函数，且最优决策规则 $\delta^*=(\tau^*,d^*)$具有如下结构：

$$\tau^* = \inf\{n \in N : \pi_n \notin (A^*, B^*)\},$$

$$d^* = \begin{cases} 1, & \pi_{\tau^*} \geqslant B^*, \\ 0, & \pi_{\tau^*} \leqslant A^*. \end{cases} \tag{6.6-6}$$

其中 A^*，B^* 按如下方式确定：

(i)若方程

$$T\rho(\pi) + c = g(\pi) \quad (\pi \in [0,1]) \tag{6.6-7}$$

在[0,1]中有解，且有两个不同的解 A^*，B^*，$A^* < B^*$，则

$$A^* < \frac{b}{a+b} < B^*;$$

(ii)如果方程(6.6-7)只有唯一解$\frac{b}{a+b}$或无解，则

$$A^* = \frac{b}{a+b} = B^*.$$

证 按推论 6.5.3，方程(6.6-2)，(6.6-3)具有唯一解 ρ. 因此，按推论 6.5.1，ρ 为 π-Bayes 最小风险函数. 由此事实及 $g(0) = 0 = g(1)$ 可得 $\rho(0) = 0 = \rho(1)$. 按引理6.6.1，$Tg(0) = 0 = Tg(1)$. 从而，$TQ^n_{(1,c)}g(0) = TQ^n_{(1,c)}g(1) = 0$. 让 $n \to \infty$，即得，$T\rho(0) = 0 = T\rho(1)$. 按引理 6.6.2，$T\rho(\pi) + c$ 与 $g(\pi)$ 同时上凸. 因此，当 $c > 0$ 时，方程(6.6-7)的不同解不可能多于二个. 若方程(6.6-7)有两个不同解 A^* 和 B^*，且 $A^* < B^*$，则必有

$$A^* < \frac{b}{a+b} \leqslant B^*.$$

(参看图 6.6-2). 注意，在(A^*, B^*)中，有 $\rho(\pi) < g(\pi)$. 故

$$\tau^* = \inf\{n \in N: \rho(\pi_n) = g(\pi_n)\}$$
$$= \inf\{n \in N: \pi_n \notin (A^*, B^*)\};$$
$$\left\{\pi_{\tau^*} \geqslant \frac{b}{a+b}\right\} = \{\pi_{\tau^*} \geqslant B^*\};$$
$$\left\{\pi_{\tau^*} < \frac{b}{a+b}\right\} = \{\pi_{\tau^*} \leqslant A^*\}.$$

如果方程(6.6-7)只有一个解或无解，则

$$\rho(\pi) \equiv g(\pi).$$

按推论 6.5.2，有

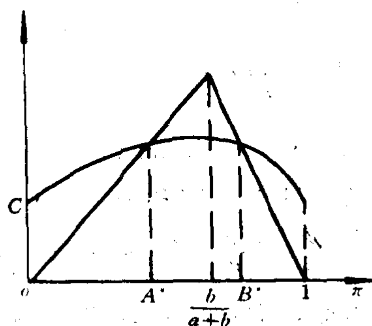


图 6.6-2　方程(6.6-7)之解 $A^*, B^*\left(A^*<\frac{b}{a+b}<B^*\right)$.

$$\tau_1^*=0,\ d^*=\begin{cases}1, & \pi_0\geqslant\frac{b}{a+b};\\ 0, & \pi_0<\frac{b}{a+b}.\end{cases}$$

因此,让 $A^*=\frac{b}{a+b}=B^*$,即得所要的结论.

从上述定理中可以看出,对于非平凡的 π-Bayes 最优决策,相应的最优决策规则 $\delta^*=(\tau^*,d^*)$ 被转化为方程(6.6-7)中确定两个不同解 A^*、B^*(即所谓阈值). 此转化的成功关键在于统计系 Ⅱ 关于 P_π 具有齐次 Markov 性. 然而,为了决定阈值 A^*,B^*,往往需要将问题转化成非 Bayes 统计系.

(Ⅰ)统计系 $\Phi=(\varphi_n,\mathscr{F}_n^\xi,n\in\mathrm{N})$,其中

$$\varphi_0=1,\varphi_n=\prod_{k=1}^{n}\varphi(\xi_k);\text{其中 }\varphi(x)=\frac{\nu_1(x)}{\nu_0(x)}.$$

$$\mathscr{F}_0^\xi=\{\varnothing,\Omega\},\mathscr{F}_n^\xi=\sigma(\xi_1,\xi_2,\cdots,\xi_n).$$

按引理 6.4.2,有

$$\pi_n^\pi=\left(1+\frac{1-\pi}{\pi\varphi_n}\right)^{-1}.\ (\pi\in[0,1],n\in\mathrm{N});$$

$$\varphi_n=\frac{1-\pi}{\pi}\frac{\pi_n^\pi}{1-\pi_n^\pi}\ (\forall\,\pi\in[0,1]).$$

假设方程(6.6-7)有两个不同解 A^*、B^*，$A^*<B^*$，则

$$\tau^* = \inf\{n \in \mathrm{N}: \pi_n \notin (A^*, B^*)\}$$
$$= \inf\left\{n \in N: \varphi_n \notin \frac{1-\pi_0}{\pi_0}(\overline{A}, \overline{B})\right\},$$

其中 $\overline{A}=\dfrac{A^*}{1-A^*}$，$\overline{B}=\dfrac{B^*}{1-B^*}$.

令 $\tau_\pi^* = \tau^*|_{\pi_0=\pi}$，则

$$\tau_\pi^* = \inf\left\{n \in \mathrm{N}: \varphi_n \notin \frac{1-\pi}{\pi}(\overline{A}, \overline{B})\right\}. \quad (\forall\, \pi \in [0,1]).$$

(Ⅱ)统计系 $\Lambda=(\lambda_n, \mathscr{F}_n^\xi, n\in\mathrm{N})$，其中

$$\lambda_0 = 0,\ \lambda_n = \sum_{k=1}^n \ln\varphi(\xi_k) = \ln\varphi_n.$$

类似地讨论立即可得

$$\tau_\pi^* = \inf\left\{n \in \mathrm{N}: \lambda_n \notin \left(\ln\frac{1-\pi}{\pi} + \ln\overline{A}, \ln\frac{1-\pi}{\pi} + \ln\overline{B}\right)\right\}$$
$$(\forall\, \pi \in [0,1]).$$

现在，我们基于统计系 Λ 考虑 Wald 意义下的最优决策问题. 设

$$\Delta(\Pi) = \{\delta = (\tau, d): \tau = \inf\{n \in \mathrm{N}: \pi_n \notin (A,B)\},$$
$$d = \begin{cases} 1, & \pi_\tau \geqslant B, \\ 0, & \pi_\tau \leqslant A, \end{cases} \quad 0 \leqslant A \leqslant B \leqslant 1\Big\}.$$

按定理 6.6.1，π-Bayes 最优决策规则 $\delta^*=(\tau^*, d^*)\in\Delta(\Pi)$. 这是在统计系 Π 下得到的. 但在统计系 Λ 下，类似地作法不一定能得到类似的结论. 这涉及到 Wald 意义下的最优决策规则的存在性问题.

定义 6.6.1 设 $\alpha\geqslant0$，$\beta\geqslant0$，且 $\alpha+\beta<1$. 记

$$\Delta(\Lambda) = \{\delta = (\tau, d): \tau = \inf\{n \in \mathrm{N}: \lambda_n \notin (A,B)\},$$
$$d = \begin{cases} 1, & \lambda_\tau \geqslant B, \\ 0, & \lambda_\tau \leqslant A. \end{cases} \quad -\infty < A \leqslant B < \infty\Big\};$$

$$\Delta^\xi(\alpha,\beta) = \{\delta = (\tau, d) \in \Delta(\Lambda): \alpha(\delta) \leqslant \alpha, \beta(\delta) < \beta,$$

$$E^i\tau < \infty (i = 0,1)\}.$$

其中 $\alpha(\delta)=P^1(d=0)$，$\beta(\delta)=P^0(d=1)$.

若∃ $\tilde{\delta}=(\tilde{\tau},\tilde{d})\in\Delta^\varepsilon(\alpha,\beta)$，∋$E^i\tilde{\tau}\leqslant E^i\tau$ ($\forall\ \delta=(\tau,d)\in\Delta^\varepsilon(\alpha,\beta),i=0,1$)，则称 $\tilde{\delta}$ 为 Wald 意义下的最优决策规则.

Wald 提出过如下问题，假如两类偏差分别控制在水平 α 和 β 下，问相应的停时 $\tilde{\tau}$ 就其平均而言是否最小，即是说，设 $\alpha\geqslant 0,\beta\geqslant 0$，且 $\alpha+\beta<1$. 假如∃ $\tilde{\delta}=(\tilde{\tau},\tilde{d})\in\Delta^\varepsilon(\alpha,\beta)$，∋$\alpha(\tilde{\delta})=\alpha,\beta(\tilde{\delta})=\beta$. 问是否有

$$E^i\tilde{\tau}\leqslant E^i\tau\ (\forall\ \delta=(\tau,d)\in\Delta^\varepsilon(\alpha,\beta),i=0,1).$$

这个问题是统计学中著名的 Wald 猜测. Lehmann 在 1959 年证明过相应的结论. 但明显只具有启发性. 最近作者研究逆 π-Bayes 决策理论，发现 Wald 问题实质上是一个逆 π-Bayes 决策问题[22]. 这里我们不介绍这方面的理论，有兴趣的读者可以追查文献[22].

§6.7 阈值的估计

假设在 $\Delta^\varepsilon(\alpha,\beta)$ 中，Wald 意义下的最优决策规则 $\tilde{\delta}=(\tilde{\tau},\tilde{d})$ 存在，则 $\alpha(\tilde{\delta})=\alpha,\beta(\tilde{\delta})=\beta$，且

$$\tilde{\tau}=\inf\{n\in\mathbf{N}:\lambda_n\notin(\tilde{A},\tilde{B})\}.$$

其中 $-\infty\leqslant\tilde{A}<0<\tilde{B}\leqslant\infty$. 一般称 $\tilde{A},\tilde{B}$ 为阈值. 要精确决定 $\tilde{A}$、$\tilde{B}$ 是困难的. 本节的目的是寻求这些阀值的有用估计. 这些估计来自于考查 $\alpha(\tilde{\delta})$、$\beta(\tilde{\delta})$ 与相应阈值 $\tilde{A},\tilde{B}$ 之间的关系.

定理 6.7.1 设 $\delta\in\Delta^\varepsilon(\alpha,\beta)$，$A<0<B$ 为相应的阈值.

若 $\quad P^i(\tau<\infty)=1\ (i=0,1)$，且 $\alpha(\delta)\vee\beta(\delta)<1$，

则 $\quad \ln\left(\dfrac{\alpha(\delta)}{1-\beta(\delta)}\right)\leqslant A<0<B\leqslant\ln\left(\dfrac{1-\alpha(\delta)}{\beta(\delta)}\right).$

证 $\alpha(\delta)=P^{1}(d=0)=P^{1}(\lambda_{\tau}\leqslant A)=\sum_{n=0}^{\infty}P^{1}(\tau=n,\lambda_{n}\leqslant A)$.

$$P^{1}(\tau=n,\lambda_{n}\leqslant A)=P^{1}(\lambda_{k}\in(A,B),1\leqslant k\leqslant n-1;\lambda_{n}\leqslant A)$$

$$=\int_{U}\prod_{k=1}^{n}\varphi(\xi_{k})P^{0}(\mathrm{d}\omega)=\int_{U}\mathrm{e}^{\lambda_{\tau}}P^{0}(\mathrm{d}\omega)$$

$$\leqslant\mathrm{e}^{A}\cdot P^{0}(U),$$

其中
$$U=\{\tau=n,\lambda_{n}\leqslant A\}$$

$$\Rightarrow\alpha(\delta)\leqslant\mathrm{e}^{A}\cdot P^{0}(d=0)=\mathrm{e}^{A}(1-\beta(\delta))$$

$$\Rightarrow\ln\left(\frac{\alpha(\delta)}{1-\beta(\delta)}\right)\leqslant A.$$

同理可证 $B\leqslant\ln\left(\frac{1-\alpha(\delta)}{\beta(\delta)}\right)$.

推论 6.7.1 如果 $\widetilde{\delta}=(\widetilde{\tau},\widetilde{d})\in\Delta^{s}(\alpha,\beta)$ 为 Wald 意义下的最优决策规则，则相应的阈值 $\widetilde{A}$、$\widetilde{B}$ 具有如下估计

$$\ln\left(\frac{\alpha}{1-\beta}\right)\leqslant\widetilde{A},\widetilde{B}\leqslant\ln\left(\frac{1-\alpha}{\beta}\right).$$

证 注意，$\alpha(\widetilde{\delta})=\alpha,\beta(\widetilde{\delta})=\beta$. 于是，在定理 6.7.1 中以 $\widetilde{\delta}$ 代替 δ，立即得到所要求的结论.

注 1 按上述推论，由 $\widetilde{A}<0$ 可知

$$\frac{\alpha}{1-\beta}<1,\Rightarrow\alpha+\beta<1.$$

此事实也可由 $\widetilde{B}>0$ 得出. 因此，$\widetilde{A}<0<\widetilde{B}$ 的存在性之必要条件是 $\alpha+\beta<1$. 若 $\alpha+\beta=1$，则 $\widetilde{A}<\widetilde{B}$ 不可能出现，且 $\widetilde{A}=0=\widetilde{B}$，这种情形相应的 $\widetilde{\tau}=0$. 可解释为最优决策规则是平凡的.

定理 6.7.2 设 A、$B\in\mathscr{R}$，且 $A<0<B$. 此外

$$P^{i}(|\ln\varphi(\xi_{1})|>0)>0,\ E^{i}\ln\varphi(\xi_{1})\neq0\ (i=0,1),$$

其中 $\varphi(x)=\frac{\nu_{1}(x)}{\nu_{0}(x)}$. 则

$$P^{i}(\tau<\infty)=1(i=0,1),$$

且
$$\exists t_{0}>0,\ni E^{i}\exp(t\tau)<\infty\ (\forall t\leqslant t_{0},i=0,1).$$

其中
$$\tau=\inf\{n\in\mathrm{N}:\lambda_{n}\notin(A,B)\}.$$

证 设 $Z_k=\lambda_k-\lambda_{k-1}(k\geqslant 1)$.

（Ⅰ）若 $C\geqslant B-A$，且 $p_i\triangleq P^i(|Z_1|\leqslant C)<1$，$p_ie^t<1$，则

$$E^i\exp(t\tau)\leqslant 1+(e^t-1)\sum_{n=0}^{\infty}(p_ie^t)^n<\infty.$$

事实上，按 τ 之定义，有

$$\begin{aligned}&(\tau\geqslant n)=\{\lambda_k\in(A,B),0\leqslant k\leqslant n-1\}\\ &\quad\subseteq\{|Z_k|\leqslant C,1\leqslant k\leqslant n-1\}\\ &\quad\Rightarrow P^i(\tau\geqslant n)\leqslant p_i^{n-1}(n\geqslant 1)\\ &\quad\Rightarrow P^i(\tau\geqslant n)\to 0 \text{ 当 } n\to\infty \text{ 时}\\ &\quad\Rightarrow P^i(\tau<\infty)=1.\ (i=0、1).\end{aligned}$$

令 $\Delta\exp(mt)=\exp(mt)-\exp((m-1)t)$，则

$$\begin{aligned}E^ie^{t\tau}&=\sum_{k=1}^{\infty}e^{kt}P^i(\tau=k)\\ &=\sum_{k=1}^{\infty}\Big(\sum_{m=1}^{k}P^i(\tau=k)\Delta e^{mt}+P^i(\tau=k)\Big)\\ &=1+\sum_{m=1}^{\infty}\Big(\sum_{k=m}^{\infty}P^i(\tau=k)\Delta e^{mt}\Big)\\ &=1+\sum_{m=1}^{\infty}P^i(\tau\geqslant m)\Delta e^{mt}\\ &\leqslant 1+\sum_{m=1}^{\infty}p_i^{m-1}\Delta e^{mt}\\ &=1+(e^t-1)\sum_{m=0}^{\infty}(p_ie^t)^m<\infty.\end{aligned}$$

（Ⅱ）若 $\exists\ C\geqslant B-A$，$\ni p_i\triangleq P^i(|Z_1|\leqslant C)<1$，则可选 $t_0>0$，使得 $p_ie^{t_0}<1$，$E^i\exp(t\tau)<\infty$ $(\forall\ t\leqslant t_0,i=0、1)$.

事实上，由结论（Ⅰ）及 $p_ie^t\leqslant p_ie^{t_0}$ 可得

$$E^i\exp(t\tau)\leqslant E^i\exp(t_0\tau)<\infty.$$

（Ⅲ）若对 $\forall\ C\geqslant B-A$，有 $P^i(|Z_1|\leqslant C)=1$，但 $\exists\ m\geqslant 1$，$\ni \overline{p}_i\triangleq P^i(|\lambda_m|\leqslant C)<1$，则当 $\overline{p}_i^{\frac{1}{m}}\cdot e^t<1$ 时，有

$$E^i\exp(t\tau) < \infty (i = 0,1).$$

从而,若选 $t_0>0$,使得 $e^{t_0}\cdot \overline{p_i}^{\frac{1}{m}}<1$,则

$$E^i\exp(t\tau) < \infty \ (\forall\ t \leqslant t_0, i = 0,1).$$

事实上,令 $\lambda_{m,k}=\sum_{i=1}^{m} Z_{(k-1)m+i}$,则有

$$\lambda_{km} = \sum_{j=1}^{k}\lambda_{m,j}.$$

由 ξ 的独立同分布性可推出

$$P^i(\tau \geqslant km) \leqslant P^i(\lambda_{jm} \in (A,B), 1 \leqslant j \leqslant k-1)$$

$$\leqslant P^i(|\lambda_{m,j}| \leqslant C, 1 \leqslant j \leqslant k-1) \leqslant \overline{p_i}^{k-1}$$

$$\Rightarrow P^i(\tau \geqslant n) \leqslant \overline{p_i}^{[\frac{n}{m}]-1} \leqslant \overline{p_i}^{-2}\,\overline{p_i}^{\frac{n}{m}} \xrightarrow[n\to\infty]{} 0$$

$$\Rightarrow E^i e^{t\tau} = 1 + \sum_{n=1}^{\infty} P^i(\tau \geqslant n)\Delta e^{nt}$$

$$\leqslant 1 + \sum_{n=1}^{n_0} P^i(\tau \geqslant n)\Delta e^{nt} + (e^t - 1)\sum_{n=n_0}^{\infty}\overline{p_i}^{\frac{n}{m}-2} e^{(n-1)t} < \infty,$$

其中 $\overline{P_i}^{\frac{1}{m}}e^t<1$,$n_0$ 满足 $\frac{n_0}{m}-2>0$.

(Ⅳ)若对 $\forall\ C\geqslant B-A$,及 $m\geqslant1$,有 $P^i(|\lambda_m|\leqslant c)=1$,则此种情形必与定理中的假设矛盾. 从而,不可能出现.

事实上,对此种情形,必 $\lambda_m(\mathrm{mod}P^i)$ 有界($\forall\ m\geqslant1$). 因此,$\frac{\lambda_m}{m} \xrightarrow[m\to\infty]{} 0(\mathrm{mod}P^i)$,按 ξ 的独立同分布性,

$$E^i\frac{\lambda_m}{m} = E^i Z_1 (\forall\ m \geqslant 1).$$

按 Lebesque 有界收敛定理,得

$$E^i Z_1 = \lim_{m\to\infty} E^i\frac{\lambda_m}{m} = 0.$$

这与 $E^iZ_1\neq0$ 之假设矛盾.

综上分析,即得本定理之结论.

注 2 当上述定理的条件满足时，对$\forall\ A<0<B$，相应的停时τ有限$(\mathrm{mod}P^i)$，且$E^i\tau<\infty$. 事实上，有

$$1+t_0E^i\tau\leqslant E^i\mathrm{e}^{t_0\tau}<\infty$$

$$\Rightarrow E^i\tau\leqslant\frac{1}{t_0}(E^i\mathrm{e}^{t_0\tau}-1)<\infty.$$

这个结果正是定理 6.7.1 所需要的条件.

第七章　附录 Markov 链

§7.1　定义与存在性

设$(\Omega,\mathscr{F},P)$为概率空间. $\mathscr{N}=\{0,1,2,\cdots\}$, $\xi=(\xi_n,n\in\mathscr{N})$为定义在$(\Omega,\mathscr{F},P)$上取值于可测空间$(E,\mathscr{E})$的$E$-值 r. v. 若$E=$ R,则ξ为 r. v. 列. 以后称E为状态空间,$(E,\mathscr{E})$为状态可测空间.

定义 7.1.1　若对$\forall\ n,m\in\mathscr{N}$,且$m\leqslant n$,有

$$P(\xi_n^{-1}(B)|\xi_0,\xi_1,\xi_2,\cdots,\xi_m)=P(\xi_n^{-1}(B)/\xi_m)\text{(a. s)}$$
$$(\forall\ B\in\mathscr{B}),$$

则称ξ为(E-值)Markov 链. 若E具有有限个元素,则称ξ为有限 Markov 链. 若E具有可数个元素,则称ξ为可列 Markov 链.

此定义中的条件有个直观解释:已知现在的状态,系统的未来状态独立它的过去状态.

定义四元函数$p(\cdot,\cdot;\cdot,\cdot)$如下:

$$p(m,x;n,B)=p(\xi_n^{-1}(B)/\xi_m=x)\ \ ((m,x;n,B)$$
$$\in \mathrm{N}\times E\times \mathrm{N}\times\mathscr{E}).$$

则此函数具有下列性质:

(i)对$\forall\ (m,n)\in\mathrm{N}\times\mathrm{N}$, $m\leqslant n$,及$x\in E$,函数$p(m,x;n,B)$为$\mathscr{E}$上的概率测度;

(ii)对$\forall\ (m,n)\in\mathrm{N}\times\mathrm{N}$, $m\leqslant n$及$B\in\mathscr{E}$,函数$p(m,x;n,B)$是$\mathscr{E}$-可测的;

(iii)对$\forall\ n\in\mathrm{N}$, $p(n,x;n,B)=1_B(x)$ $(\forall\ (x,B)\in E\times\mathscr{E})$;

(iv)Chapman-Kolmogorov 方程(以后简称 Ch. -K. 方程)成立:

$$p(m,x;n,B)=\int_E p(m,x;u,\mathrm{d}y)p(u,y;n,B),$$

其中$(m,n)\in \mathrm{N}\times\mathrm{N},m\leqslant u\leqslant n,u\in\mathrm{N},(x,B)\in E\times\mathscr{E}$.

(事实上,按条件期望之性质,有

$$\begin{aligned}P(\xi_n^{-1}(B)/\xi_m)&=E(1_{\xi_n^{-1}(B)}/\xi_m)\\&=E(E(1_{\xi_n^{-1}(B)}/\xi_m,\xi_u)/\xi_m)\\&=E(E(1_{\xi_n^{-1}(B)}/\xi_u)/\xi_m)\\&=\int_E p(u,y;n,B)p(m,\xi_m;u,\mathrm{d}y).\end{aligned}$$

让$\xi_m=x$即得 Ch. -K. 等式.)

定义 7.1.2 若定义在$\mathrm{N}\times E\times\mathrm{N}\times\mathscr{E}$上的四元函数$p$满足上述条件(i)～(iv),则称此四元函数为转移概率函数.

引理 7.1.1 设ξ为$(\Omega,\mathscr{F},P)$上的E-值 Markov 链具有初始分布π:$\pi(B)=P(\xi_0^{-1}(B))(\forall B\in\mathscr{E})$及转移函数$p(m,x;n,B)$,则$\xi$的有限维概率分布族按如下方式由$\pi$和$p$决定:

$$P(\xi_k^{-1}(B_k),0\leqslant k\leqslant n)=\int_{B_0}\pi(\mathrm{d}x_0)\int_{B_1}p(0,x_0;1,\mathrm{d}x_1)\cdots\int_{B_n}p(n-1,x_{n-1};n,\mathrm{d}x_n),$$

其中$B_k\in\epsilon,0\leqslant k\leqslant n,n\in\mathrm{N}$.

证 事实上,令$A_n=\bigcap_{k=0}^{n}\xi_k^{-1}(B_k)$,则利用条件期望的性质,可得

$$\begin{aligned}E(1_{A_n})&=E(E(1_{A_n}/\xi_0,\xi_1,\cdots,\xi_{n-1}))\\&=E(1_{A_{n-1}}\cdot E(1_{\xi_n^{-1}(B_n)}/\xi_0,\xi_1,\cdots,\xi_{n-1}))\\&=E(1_{A_{n-1}}\cdot E(1_{\xi_n^{-1}(B_n)}/\xi_{n-1}))\\&=E(1_{A_{n-2}}\cdot E(\eta_{n-1}(\xi_{n-1})/\xi_{n-2}))\\&=\cdots=E(1_{A_0}\cdot E(\eta_1(\xi_1)/\xi_0))\\&=E(1_{B_0}(\xi_0)\cdot E(\eta_1(\xi_1)/\xi_0)),\end{aligned}$$

其中 $\eta_k(\xi_k)=1_{B_k}(\xi_k)\cdot E(\eta_{k+1}(\xi_{k+1})/\xi_k)$ $(1\leqslant k\leqslant n-1)$,

$$\eta_n(\xi_n)=1_{B_n}(\xi_n)$$

$$\Rightarrow E(1_{A_n})=\int_{B_0}\pi(\mathrm{d}x_0)\int_E\eta_1(x_1)p(0,x_0;1,\mathrm{d}x_2),$$

$$\eta_1(x_1)=1_{B_1}(x_1)\cdot\int_E\eta_2(x_2)p(1,x_1;2,\mathrm{d}x_2),$$

$$\eta_k(x_k)=1_{B_k}(x_k)\int_E\eta_{k+1}(x_{k+1})p(k,x_k;k+1,\mathrm{d}x_{k+1})$$

$$(1\leqslant k\leqslant n-1),$$

$$\eta_n(x_n)=1_{B_n}(x_n).$$

通过逐步代入即可得引理中的公式.

定理 7.1.1 设 π 为 $(E,\mathscr{E})$ 上的一个概率测度. p 为转移概率函数. 则∃ 一个概率空间 $(\Omega,\mathscr{F},P)$ 及一个定义于其上的 Markov 链 ξ, 〕ξ 具有初始分布 π 和转移概率 p.

证 取 $(\Omega_n,\mathscr{F}_n)=(E,\mathscr{E}),(\Omega,\mathscr{F})=\bigtimes_{n=0}^{\infty}(\Omega_n,\mathscr{F}_n)$. 注意, 转移概率函数 p 满足 I.-T. 定理(参看第 1 章)中的要求. 概率测度 P 按引理 7.1.1 中的公式定义, 则按 I.-T. 定理, 可在 $(\Omega,\mathscr{F},P)$ 上定义一个 E-值 r.v. 列 ξ, 使得 ξ 的有限维分布族具有引理 7.1.1 中的结构; 现在, 来证明: 此 ξ 为 Markov 链.

首先证明 ξ 具有转移函数 p. 设 $B_0\in\sigma(\xi_0)$, 则

$$E(1_{B_0}(\xi_0)1_{B_1}(\xi_1))=\int_{B_0}\pi(\mathrm{d}x_0)\int_{B_1}p(0,x_0;1,\mathrm{d}x_1)$$

$$=E(1_{B_0}(\xi_0)\cdot p(0,\xi_0;1,B_1))$$

$$\Rightarrow P(\xi_1^{-1}(B_1)/\xi_0)=p(0,\xi_0;1,B_1).$$

设 $B_n\in\sigma(\xi_n),\mu_n(B)=P(\xi_n^{-1}(B))=P(E^{n-1}\times\xi_n^{-1}(B))$, 则

$$E(1_{B_n}(\xi_n)\cdot 1_{B_{n+1}}(\xi_{n+1}))=E(1_{E^{n-1}}\cdot 1_{B_n}(\xi_n)\cdot 1_{B_{n+1}}(\xi_{n+1}))$$

$$=E(1_{E^{n-1}}1_{B_n}(\xi_n)\cdot p(n,\xi_n;n+1,B_{n+1}))$$

$$=E(1_{B_n}(\xi_n)\cdot p(n,\xi_n;n+1,B_{n+1})).$$

$$\Rightarrow P(\xi_{n+1}^{-1}(B_{n+1})/\xi_n) = p(n,\xi_n;n+1,B_{n+1}).$$

类似地推理，即可证明：

$$P(\xi_n^{-1}(B_n)/\xi_m) = p(m,\xi_m;n,B_n)\ (m \leqslant n).$$

将 ξ_m 改成 x，由上式立即得到 p 为 ξ 的转移函数的结论.

现在证明：ξ 具有 Markov 性. 这等价于证明：

$$E1_{D^n}(\xi_0,\xi_1,\cdots,\xi_{n+1}) = E(1_{D^{n-1}}(\xi_0,\xi_1,\cdots,\xi_n)E(1_{B^{n+1}}(\xi_{n+1})/\xi_n)).$$

其中 $D^n = \bigtimes_{k=0}^{n+1} B_k, B_k \in \varepsilon (0 \leqslant k \leqslant n+1), n \in \mathrm{N}$.

按 P 的构造，有

$$\begin{aligned} E1_{D^n}(\xi_0,\xi_1,\cdots,\xi_{n+1}) &= E(1_{D^{n-1}}(\xi_0,\xi_1,\cdots,\xi_n) \\ &\quad\cdot p(n,\xi_n;n+1,B_{n+1})) \\ &= E(1_{D^{n-1}}(\xi_0,\xi_1,\cdots,\xi_n) \\ &\quad\cdot E(1_{B_{n+1}}(\xi_{n+1}/\xi_n)). \end{aligned}$$

故 ξ 具有 Markov 性.

注 1 如果让 $(E,\varepsilon)=(\mathrm{R},\mathscr{B})$，则可引用 Kolmogorov 定理(参看第 1 章)得到所要的结论. 但对于一般的可测空间 $(E,\mathscr{E})$，宜于采用 I. -T. 定理，因为它不需要 $(E,\mathscr{E})$ 上的某种拓扑结构.

注 2 上述定理解决 Markov 链的存在问题. 自然还有一个唯一性问题需要考虑. 有人举了一个例(参看[20])证实唯一性不能保证. 这里问题出在 Ch. -K. 方程上. 假定 $\{(\Omega,\mathscr{F},P);\xi\}$ 为一个 E-值 r. v. 列. 为使 ξ 的转移函数满足 Ch. -K. 方程，并不要求 ξ 必须具有 Markov 性. 比如，可设 ξ 满足要求：

$$P(\xi_n \in B_n/\xi_{m_1},\xi_{m_2}) = P(\xi_n \in B_n/\xi_{m_2}),$$

其中 $m_1,m_2,n \in N, m_1 < m_2 \leqslant n, B_n \in \varepsilon$.

则 $p(m,x;n,B)=P(\xi_n^{-1}(B)/\xi_m)$ 满足 Ch. -K. 方程. 但此 ξ 一般不是这里定义的 Markov 链. 于是在 Markov 链理论中，多阶 Markov 链是一个重要问题.

定义 7.1.3 若转移函数 $p(m+k,x;n+k,B)$ 与 k 无关，则称此转移函数是齐次的(或齐时的). 记

$$p(k,x;B)=p(m,x;m+k,B).$$

称 $p(k,x;B)$ 为 k 步转移函数.

按 Ch.-K. 方程，k 步转移函数可以通过一步转移函数 $p(1,x;B)=p(x,B)$ 表示出来：

$$p(k,x;B)=\int_E p(x,dy)p(k-1,y;B)$$

$$=\int_{E^{k-1}} p(x;dy_1)p(y_1;dy_2)\cdots p(y_{k-2};dy_{k-1})p(y_{k-1};B).$$

定理 7.1.2 设 $\{(\Omega,\mathscr{F},P);\xi\}$ 为 E-值 Markov 链. η_0 是 $(\mathrm{N},\mathscr{N})\to(\mathrm{N},\mathscr{N})$ 的可测映射，其中 $\mathscr{N}$ 表示 N 中子集之全体. Q 为 $(\mathrm{N},\mathscr{N})$ 上的概率测度. $\tilde{P}=Q\times P$ 为 $\mathscr{N}\times\mathscr{F}$ 上的概率测度. $r_n(u,\omega)=(n+\eta_0(u),\xi(n+\eta_0(u),\omega))$，其中 $n\in\mathscr{N}$，$(u,\omega)\in\mathrm{N}\times\Omega$，$\xi(n,\omega)=\xi_n(\omega)$. 则 $r=(r_n,n\in\mathscr{N})$ 为概率空间 $(\mathrm{N}\times\Omega,\mathscr{N}\times\mathscr{F},\tilde{P})$ 上的齐次 Markov 链.

证 设 $\Gamma=\Delta\times B\in\mathscr{N}\times\mathscr{E}$，$r_n=(r'_n,r''_n)$，则

$$\tilde{P}(r_{n+1}\in\Gamma/r_n)=\tilde{P}(r'_n+1\in\Delta,\xi(r'_n+1,\omega)\in B/r_n)$$

$$=1_\Delta(r'_n+1)\tilde{P}(\xi(r'_n+1,\omega)\in B/r_n)$$

$$=1_\Delta(r'_n+1)\cdot P(\xi(k+1,\omega)\in B/\xi(k,\omega))|_{k=r'_n}$$

$$=1_\Delta(r'_n+1)\cdot p(k,\xi_k;k+1,B)|_{k=r'_n}.$$

其中 $p(k,x;k+1,B)$ 为 ξ 的转移函数. 令 $z=(z',z'')\in\mathrm{N}\times E$，则由上式得

$$\tilde{p}(n,Z;n+1,\Gamma)\triangleq\tilde{P}(r_{n+1}\in\Gamma/r_n)|_{r_n=Z}$$

$$=1_\Delta(Z'+1)\cdot p(Z',Z'';Z'+1,B).$$

此式的右边与 n 无关，故 r 的转移函数 $\tilde{p}$ 是齐次的.

让 $\Gamma=\Delta\times B\in\mathrm{N}\times\mathscr{E}$，则

$$\tilde{P}(r_{n+1}\in\Gamma/r_0,r_1,\cdots,r_n)=\tilde{P}(r'_n+1\in\Delta,\xi(r'_n+1,\omega)\in B/(r_0,r_1,\cdots,r_n))$$

$$=1_\Delta(r'_n+1)\cdot P(\xi_{k+1}\in B/\xi_0,\xi_1,\cdots,\xi_k)|_{k=r'_n}$$

$$=1_\Delta(r'_n+1)P(\xi_{k+1}\in B/\xi_k)|_{k=r'_n}$$

$$= \widetilde{P}(r_{n+1} \in \Gamma / r_n).$$

故 r 具有 Markov 性.

注 3 上述定理中的 η_0 和 Q 的设计具有很大的任意性. 下面的推论就是这种任意性的一个特例.

推论 7.1.1 设 ξ 为 $(\Omega, \mathscr{F}, P)$ 上的 E-值 Markov 链. $Z_n^* = (n, \xi_n)$，则 $Z^* = (Z_n^*, n \in \mathrm{N})$ 为概率空间 $(\mathrm{N} \times \Omega, \mathscr{N} \times \mathscr{E}, P^*)$ 上的齐次 Markov 链，其中

$$P^* = \Lambda^* \times P, \lambda^*(\Delta) = \begin{cases} 1, & 0 \in \Delta, \\ 0, & 0 \notin \Delta. \end{cases} \quad (\forall\ \Delta \in \mathrm{N})$$

且转移函数为

$$p^*(z; \Gamma) = 1_\Delta(z' + 1) \cdot p(z', z''; z' + 1, B),$$

其中 $z = (z', z'') \in \mathrm{N} \times E, \Gamma = \Delta \times B \in \mathrm{N} \times \mathscr{E}$.

证 让 $Q(\eta_0 \in \Delta) = \lambda^*(\Delta), \eta_0 = 0$，则 $r_n = z_n^*$. 从而，推论变为定理 7.1.2 的结论.

注 4 定理 7.1.2 和推论 7.1.1 表明：多维齐次 Markov 链的分量可以不是齐次的. 此外，由于这种齐次性转化的可能，致使人们将研究的目标集中在齐次 Markov 链上.

后面几节将介绍齐次可列或齐次有限 Markov 链. 这里首先定义一些符号, 设状态空间 E 是有限的或可列的. $\{(\Omega, \mathscr{F}, P); \xi\}$ 为 E-值齐次 Markov 链. 则有关常用符号定义如下：

$$p_{ij} = P(\xi_{n+1} = j / \xi_n = i) = P(\xi_1 = j / \xi_0 = i) \ (n \in \mathrm{N}, i、j \in E);$$

$$p_{ij}^{(k)} = P(\xi_{n+k} = j / \xi_n = i) = P(\xi_k = j / \xi_0 = i) \ (n, k \in \mathrm{N}, i、j \in E).$$

$$p_{ij}^{(0)} = \begin{cases} 1, & j = i, \\ 0, & j \neq i. \end{cases} \quad (i、j \in E).$$

$$p_{ij}^{(1)} = p_{ij}.$$

$$\mathscr{P} = [p_{ij}]_{i,j \in E}. \quad \mathscr{P}^{(k)} = [p_{ij}^{(k)}]_{i,j \in E} (\forall\ k \in \mathrm{N}).$$

上述符号解释如下：$p_{ij}^{(k)}$ 表示系统在起始时刻从状态 i 出发经 k 步到达状态 j 的概率，又称它为 k 步转移概率. 按 Ch.-K. 方程，它可

以用一步转移概率表达出来：

$$p_{ij}^{(k)} = \sum_{\substack{i_n \in E \\ 1 \leqslant n \leqslant k-1}} p_{ii_1} p_{i_1 i_2} \cdots p_{i_{k-1} j}.$$

$\mathrm{P}^{(k)}$可解释为 k 步转移矩阵，具有下列特点：

(i)$\mathrm{P}^{(k)}$的元素非负，每行元素之和为 1，即

$$\sum_{j \in E} p_{ij}^{(k)} = 1 \ (\forall\, i \in E);$$

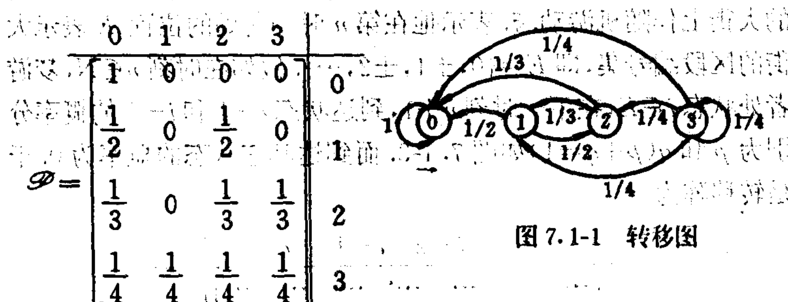
(ii)$\mathrm{P}^{(0)} = I$ ($|E| \times |E|$阶单位阵，$|E|$表 E 中元素个数)；

(iii)$\mathrm{P}^{(k+l)} = \mathrm{P}^{(k)} \cdot \mathrm{P}^{(l)} = \mathrm{P}^{k+l}$(这由 Ch.-K. 方程推出).

定义 7.1.4 若矩阵族($\mathscr{P}^{(k)}, k \in \mathrm{N}$)具有性质(i)～(iii)，则称 $\mathscr{P}^{(k)}$ 为转移阵($\forall\, k \in \mathrm{N}$). 按性质(iii)，$\mathscr{P}^{(k)}$ 由 P 决定，以后讲到转移阵就是指 P(即一步转移阵).

按定理 7.1.1，E-值齐次 Markov 链 ξ 由两个因素决定：一是初始分布 π，一是转移阵 P. 因此，这类 Markov 链讨论的中心问题是：P 的结构，π 的设计及 π 与 P 之间的关系.

在 Markov 链理论中人们常常将转移阵 $\mathscr{P}$ 用一种所谓转移图(图 7.1-1)表示出来，以求得直观，例如：设 $E = \{0,1,2,3\}$.



$$\mathscr{P} = \begin{array}{c|c} \begin{array}{cccc} 0 & 1 & 2 & 3 \end{array} & \\ \hline \begin{bmatrix} 1 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ \frac{1}{3} & 0 & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{bmatrix} & \begin{array}{c} 0 \\ 1 \\ 2 \\ 3 \end{array} \end{array}$$

图 7.1-1 转移图

例 7.1.1 考虑某个按顺序执行的"任务"序列. 状态空间 $E = \mathrm{N}$. 假定第 n 个"任务"成功之后向一"任务"转移的概率为 p_n. 失败之后就转回到开始的"任务". 其概率为 q_n，$p_n + q_n = 1$. 这个模型

的转移阵为

$$
\begin{array}{c|c}
\begin{array}{ccccc} 0 & 1 & 2 & & \cdots \end{array} & \\
\hline
P=\begin{bmatrix} q_1 & p_1 & 0 & 0 & \cdots \\ q_2 & 0 & p_2 & 0 & \cdots \\ q_3 & 0 & 0 & p_2 & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \end{bmatrix} & \begin{array}{c} 0 \\ 1 \\ 2 \\ \vdots \end{array}
\end{array}
$$

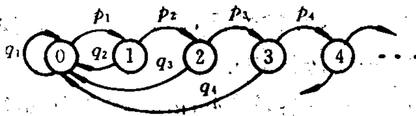


图 7.1-2　转移图

相应的链为齐次 Markov 链(在[13]中称此例为基本例)如图 7.1-2.

例 7.1.2　无限随机游动,考虑一个梦游者在一条两端无禁的大街上作随机游动. ξ_n 表示他在第 n 时刻所处的位置. E 表示大街的区段编号集,即 $E=\{0,\pm1,\pm2,\cdots\}$. 假如在时刻 $n\in \mathrm{N}$,梦游者处状态 $i\in E$,在下一时刻 $n+1$,到达状态 $i+1$ 和 $i-1$ 的概率分别为 p 和 $q(p+q=1)$ 如图 7.1-3. 而到达其它状态的概率为 0. 于是转移阵为

$$
\begin{array}{c|c}
\begin{array}{ccccccc} & & i-2 & i-1 & i & & \end{array} & \\
\hline
P=\begin{bmatrix} \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & q & 0 & p & 0 & \cdots & \cdots \\ \cdots & 0 & q & 0 & p & \cdots & \cdots \\ \cdots & 0 & 0 & q & 0 & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \end{bmatrix} & \begin{array}{c} \\ i-1 \\ i \\ i+1 \\ \\ \end{array}
\end{array}
$$

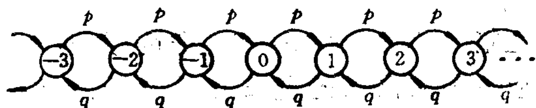


图 7.1-3　转移图

在此模型中，梦游者的游动只与 n 时刻的状态有关. 而与 n 时刻之前的状态无关. 转移阵与时刻 n 无关. 因此，此模型为一个齐次 Markov 链. 其转移阵 $P=[p_{ij}]_{i,j\in E}$ 中的元素 p_{ij} 有如下表示：

$$p_{ij}=\begin{cases}p, j=i+1,\\ q, j=i-1,\\ 0, \text{其它}.\end{cases}\quad (i,j\in E).$$

k 步转移概率 $p_{ij}^{(k)}$ 可按如下公式计算：

$$p_{ij}^{(k)}=\begin{cases}C_k^{\frac{1}{2}(k+j-i)}\cdot p^{\frac{1}{2}(k+j-i)}q^{\frac{1}{2}(k-j+i)}, & \text{当 } k+j-i \text{ 为偶数，}\\ & 0\leqslant k+j-i\leqslant 2k \text{ 时，}\\ 0, & \text{否则.}\end{cases}$$

为了证明这个公式，设 $Y=(Y_l, l\geqslant 1)$ 为取值于 $\{-1,1\}$ 的独立同分布的 r.v. 列，且

$$P(Y_l=1)=p,\ P(Y_l=-1)=q\quad (\forall\, l\geqslant 1).$$

显然，有

$$\xi_k=\xi_{k-1}+Y_k=\xi_0+\sum_{l=1}^{k}Y_l(\forall\, k\geqslant 1).$$

现设 $\xi_0=i$. 样本集 $\{Y_1,Y_2,\cdots,Y_k\}$ 的样本中有 k_1 个分量取 1，k_2 个分量取 -1. 固定 k_1 时该族相应的样本数为 $C_k^{k_1}$（即有 k_1 个分量为 1 的组合数）. 每个样本出现的概率为 $p^{k_1}\cdot q^{k_2}$. 现设在第 k 步到达状态 j，则 $\sum_{l=1}^{k}Y_l=j-\xi_0=j-i$. 由此得一方程组：

$$k_1+k_2=k,\ k_1-k_2=j-i.$$

解之得：$k_1=\frac{1}{2}(k+j-i)$，$k_2=\frac{1}{2}(k-j+i)$. 按这里的需要，k_1 和 k_2 均只能是正整数. 因此，当 $k+j-i$ 为非负偶数时，此解有意义. 其它情形相应的概率为 0. 综合起来，便得所要的公式.

作为特例，有

$$p_{ii}^{(k)}=\begin{cases}C_k^{\frac{1}{2}k}\cdot(p\cdot q)^{\frac{1}{2}k}, & \text{当 } k\in \mathrm{N}\text{，且为偶数时，}\\ 0, & \text{否则.}\end{cases}$$

$$p_{ii}^{(2n)}=C_{2n}^{n}(pq)^n=\frac{(2n)!}{(n!)^2}(pq)^n\approx\frac{(4pq)^n}{\sqrt{n\pi}}\ (\forall n\in \mathrm{N}).$$

这里最后的近似关系是由 Stirling 公式得出的：$n!\approx\left(\frac{n}{e}\right)^n\sqrt{2n\pi}$.

例 7.1.3 带有两个吸收壁的随机游动. 假如有一条街处于梦游者的房屋与湖塘之间. 当他在街道上随机游动时最终要么进屋，要么落入湖塘. 设状态空间 $E=\{0,1,\cdots,n\}$. “0”表湖塘，“n”表房屋. 让 $p_{00}=1=p_{nn}$ 分别表示落入湖塘和进入房屋如图 7.1-4（终止游动的两个可能结局）. 他从状态 i 到达状态 j 的转移概率为

$$p_{ij}=\begin{cases}p, & j=i+1,\\ q, & j=i-1,\ (i,j\in E, i\neq 0\text{ 或 }n), p+q=1.\\ 0, & \text{其它}\end{cases}$$

于是，转移阵为

$$\mathrm{P}=\begin{array}{c} \begin{array}{ccccccccc} 0 & 1 & 2 & 3 & & & n-2 & n-1 & n \end{array} \\ \begin{bmatrix} 1 & 0 & 0 & 0 & \cdots & \cdots & 0 & 0 & 0\\ q & 0 & p & 0 & \cdots & \cdots & 0 & 0 & 0\\ 0 & q & 0 & p & \cdots & \cdots & 0 & 0 & 0\\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots\\ 0 & 0 & 0 & 0 & \cdots & \cdots & q & 0 & p\\ 0 & 0 & 0 & 0 & \cdots & \cdots & 0 & 0 & 1 \end{bmatrix} \end{array} \begin{array}{c} 0\\ 1\\ 2\\ \\ n-1\\ n. \end{array}$$

此模型为齐次 Markov 链. 它的一个特点是梦游者一旦进入状态 0

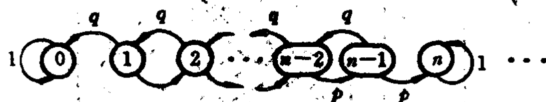


图 7.1-4　转移图

或 n，则永远也不会离开. 这种状态称为吸收壁.

§7.2　状态分类

设$\{(\Omega,\mathscr{F},P);\xi\}$为 E-值齐次 Markov 链，其中状态空间 E 有限或可数. 本节讨论系统中各种状态的特点.

定义 7.2.1　若状态 $i\in E$ 具有特点：$\exists\ m\geqslant 1, j\in E, \ni p_{ij}^{(m)}>0$；对$\forall\ n\geqslant 1, j\in E$，有 $p_{ji}^{(n)}=0$，则称 i 为非本质状态. 不具有这种性质的状态称为本质状态.

例 7.2.1　设 $E=\{0,1,2,3,4\}$. 转移阵和转移图如图 7.2-1.

$$
\begin{array}{c|c}
\begin{matrix} \quad 0 & \ 1 & \ 2 & \ 3 & \ 4 \end{matrix} & \\
\hline
P=\begin{bmatrix}
\frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 \\
0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\
\frac{1}{3} & \frac{1}{3} & 0 & \frac{1}{3} & 0 \\
\frac{1}{4} & \frac{1}{4} & 0 & \frac{1}{4} & \frac{1}{4}
\end{bmatrix} & \begin{matrix} 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{matrix}
\end{array}
$$

引理 7.2.1　若 i 为非本质状态，则划去 $\mathscr{P}$ 中相应于 i 的行和列所构成的 $\mathscr{P}_i$ 仍为转移阵.

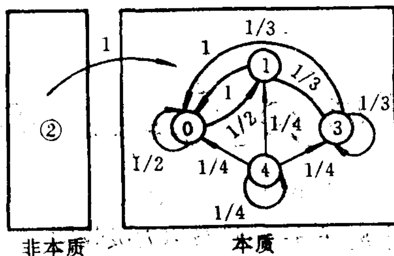


图 7.2-1 本质、非本质状态

证 事实上，按定义 7.2.1，$P^{(k)}$ 中相应于 i 的那一列元素全为 0. 因此，$P_i^{(k)}$ 满足定义 1.4 中的条件(i)($\forall\, k\geqslant 1$). 由 $P^{(k)}=P^k$ 可知：$P_i^{(k)}=P_i^k$. 于是，$P_i^{(k)}$ 满足该定义中的条件(iii)，而条件(ii)自然满足. 故 P_i 为转移阵.

按定义 7.2.1，当系统处于非本质状态时，最终必然转移到本质状态，而系统处于本质状态时，永远也不能转移到非本质状态. 基于这一缘故，人们只对本质状态感兴趣. 以后，总假定状态空间 E 中不含非本质状态.

定义 7.2.2 若状态 $i、j\in E$ 具有性质：$\exists\, m\in \mathrm{N}, m\neq 0$，使 $p_{ij}^{(m)}>0$，则称 i 可达 j. 记为 $i\to j$. 若 $i\to j$ 且 $j\to i$，则称 i 与 j 互通. 记为 $i\longleftrightarrow j$.

引理 7.2.2 互通关系"$\longleftrightarrow$"为一个等价关系，即具有自反性，对称性，传递性.

证 自返性和对称性显然. 下证传递性. 为此，设 $i\longleftrightarrow u$，$u\longleftrightarrow j$，则按定义 7.2.2，$\exists\, n\geqslant 1, m\geqslant 1$，使 $p_{iu}^{(n)}>0$，$p_{uj}^{(m)}>0$. 于是，按 Ch.-K. 方程，有

$$p_{ij}^{(n+m)}=\sum_{v\in E}p_{iv}^{(n)}p_{vj}^{(m)}\geqslant p_{iu}^{(n)}p_{uj}^{(m)}>0$$

$$\Rightarrow i\to j.$$

同理可证：$j\to i$，故 $i\longleftrightarrow j$.

注 1 从此引理之证明中可以看出：可达关系"$\to$"具有传递

性.

定义 7.2.3 设$\mathscr{F}_n=\sigma(\xi_0,\xi_1,\cdots,\xi_n)$. $(\mathscr{F}_0)=(\mathscr{F}_n,n\in\mathrm{N})$.

$$\sigma_i=\inf\{n\geqslant 1:\xi_n=i\},\ \inf\{\varnothing\}=\infty.$$

则称σ_i为系统ξ访问状态i的首达时,有时也称停时或$(\mathscr{F}_0)$-时或 Markov 时.

首达时σ_i具有下列特点:

(i) $(\sigma_i=n)=\{\xi_l\neq i,1\leqslant l\leqslant n-1;\xi_n=i\}$;

(ii) $(\sigma_i=n)\cap(\sigma_j=m)=\varnothing(n\neq m)$;

(iii) $(\sigma_i=n)\in\mathscr{F}_n(\forall\ n\in\mathrm{N})$;

(iv) $\mathscr{F}_{\sigma_i}=\{A\in\mathscr{F}:A\cap(\sigma_i=n)\in\mathscr{F}_n(\forall\ n\geqslant 1)\}$, σ_i是$\mathscr{F}_{\sigma_i}$-可测的.

现在定义与首达时σ_j有关的量.

$$f_{ij}^{(0)}=0;\ f_{ij}^{(n)}=P(\sigma_j=n/\xi_0=i)\ (\forall\ n\geqslant 1,i,j\in E);$$

$$f_{ij}=P(\sigma_j<\infty/\xi_0=i)=\sum_{n=1}^{\infty}f_{ij}^{(n)}(\forall\ i,j\in E).$$

$$\mu_j=E(\sigma_j 1_{(\sigma_j<\infty)}/\xi_0=j).$$

这些符号有如下解释:$f_{ij}^{(n)}$表示"系统从状态i出发在n时刻之前不访问状态j而在时刻n访问状态j"的概率. f_{ij}表示"系统从状态i出发最终访问状态j"的概率, μ_j表示"系统从状态j出发最终在有限时间内返回状态j的平均时".

引理 7.2.3

(i) $p_{ij}^{(n)}=\sum_{k=1}^{n}f_{ij}^{(k)}p_{jj}^{(n-k)}(\forall\ n\geqslant 1,i、j\in E)$;

(ii) $f_{ij}^{(n)}=\sum_{\substack{i_l\in E\\ i_l\neq j,1\leqslant l\leqslant n-1}}p_{ii_1}p_{i_1i_2}\cdots p_{i_{n-1}j}(\forall\ n\geqslant 1,i、j\in E)$.

证 按σ_j之定义,有

$$\{\xi_0=i,\xi_n=j\}=\{\xi_0=i\}\cap\{\xi_n=j,\sigma_j\leqslant n\}$$
$$=\bigcup_{l=1}^{n}\{\xi_l=i\}\cap\{\xi_n=j,\sigma_j=l\}.$$

利用 Markov 性，可得

$$
\begin{aligned}
p_{ij}^{(n)} &= P(\xi_n = j/\xi_0 = i) = P(\xi_0 = i, \xi_n = j/\xi_0 = i) \\
&= P\Big(\bigcup_{l=1}^{n}\{\xi_0 = i, \xi_n = j, \sigma_j = l\}/\xi_0 = i\Big) \\
&= \sum_{l=1}^{n} P(\xi_n = j, \sigma_j = l/\xi_0 = i) \\
&= \sum_{l=1}^{n} P(\xi_n = j/\xi_0 = i, \sigma_j = l) \cdot P(\sigma_j = l/\xi_0 = i) \\
&= \sum_{l=1}^{n} f_{ij}^{(l)} \cdot P(\xi_n = j/\xi_l = j) \\
&= \sum_{l=1}^{n} f_{ij}^{(l)} \cdot p_{jj}^{(n-l)}.
\end{aligned}
$$

$$
\begin{aligned}
f_{ij}^{(n)} = P(\sigma_j = n/\xi_0 = i) &= P(\xi_0 = i, \xi_k \neq j, 1 \leqslant k \leqslant n-1; \\
&\quad \xi_n = j/\xi_0 = i) \\
&= \sum_{i_{n-1}\neq j} P(\xi_k \neq j, 1 \leqslant k \leqslant n-2; \xi_{n-1} = i_{n-1}; \xi_n = j/\xi_0 = i) \\
&= \sum_{i_{n-1}\neq j} p_{i_{n-1}j} \cdot P(\xi_k \neq j, 1 \leqslant k \leqslant n-2; \xi_{n-1} = i_{n-1}/\xi_0 = i) \\
&= \sum_{i_{n-1}\neq j} \cdots \sum_{i_1\neq j} p_{i_{n-1}j} p_{i_{n-2}i_{n-1}} \cdots p_{ii_1}.
\end{aligned}
$$

引理 7.2.4 “$i \to j$”$\Leftrightarrow$“$f_{ij} > 0$”.

证 $f_{ij} > 0 \Rightarrow \exists\ m \geqslant 1, \ni f_{ij}^{(m)} > 0$. 按引理 7.2.3，有

$$p_{ij}^{(m)} \geqslant f_{ij}^{(m)} > 0 \Rightarrow i \to j \Rightarrow \exists\ n \geqslant 1, \ni p_{ij}^{(n)} > 0.$$

按引理 7.2.3，$\exists\ 1 \leqslant l \leqslant n, \ni f_{ij}^{(l)} > 0 \Rightarrow f_{ij} > 0$.

引理 7.2.5 设$(b_n, n \geqslant 1)$为非负数列，且

$$\lim_{n\to\infty} \frac{b_n}{\sum_1^n b_k} = 0,$$

则对任意收敛数列$(c_n, n \geqslant 1)$，有

$$\lim_{N\to\infty} \sum_{n=1}^{N} b_n c_{N-n} \Big(\sum_{k=1}^{n} b_k\Big)^{-1} = \lim_{m\to\infty} c_m = c^*.$$

证　令 $b_N^* = \sum_{k=1}^{N} b_k$，则对∀ $1 \leqslant n \leqslant N$，有

$$\left|\sum_{l=1}^{N} \frac{b_l c_{N-l}}{b_N^*} - c^*\right| \leqslant \frac{1}{b_N^*} \sum_{l=N-n}^{N} b_l |c_{N-l} - c^*|$$

$$+ \frac{1}{b_N^*} \sum_{l=1}^{N-n} b_l \cdot |c_{N-l} - c^*|$$

$$\leqslant \sup_k |c_k - c^*| \cdot (n+1) \sup_{0 \leqslant l \leqslant n} \left(\frac{b_{N-n+l}}{b_N^*}\right)$$

$$+ \sup_{1 \leqslant l \leqslant N-n} |c_{N-l} - c^*|$$

$$\leqslant (n+1) \sup_k |c_k - c^*| \sup_{m \geqslant N-n} \left(\frac{b_m}{b_m^*}\right)$$

$$+ \sup_{m \geqslant n} (|c_m - c^*|).$$

按假设，对∀ $\varepsilon > 0$，∃ $n \geqslant 1$，∋

$$\sup_{m \geqslant n} (|c_m - c^*|) < \frac{1}{2}\varepsilon.$$

∃ $N_0 = N_0(\varepsilon, n)$，使得 当 $N \geqslant N_0$ 时，有

$$(n+1) \sup_k (|c_k - c^*|) \cdot \sup_{m \geqslant N-n} \left(\frac{b_m}{b_m^*}\right) < \frac{1}{2}\varepsilon$$

$$\Rightarrow \left|\sum_{l=1}^{N} \frac{b_l}{b_l^*} c_{N-l} - c^*\right| \leqslant \varepsilon \text{ 当 } N \geqslant N_0 \text{ 时.}$$

引理 7.2.6　“$f_{ii} = (<) 1$”$\Longleftrightarrow$“$\sum_{n=1}^{\infty} p_{ii}^{(n)} = (<) \infty$”.

证　令 $c_n = \sum_{k=0}^{n} p_{ii}^{(k)}$，$b_n = f_{ii}^{(n)}$，$b_n^* = \sum_{k=1}^{n} b_k$，则按引理 7.2.3(i)，有

$$\sum_{k=1}^{N} p_{ii}^{(k)} = \sum_{l=1}^{N} b_l c_{N-l}, \quad c_N = 1 + \sum_{l=1}^{N} b_l \cdot c_{N-l}$$

$$\Rightarrow \frac{c_N - 1}{c_N} \leqslant \sum_{l=1}^{N} b_l \leqslant f_{ii} (\forall\, \mathrm{N} \geqslant 1).$$

若 $c_N \xrightarrow[N \to \infty]{} \infty$，则 $1 \leqslant f_{ii} \Rightarrow f_{ii} = 1$.

若$\lim\limits_{N\to\infty} c_N<\infty$，则按引理 7.2.5，有

$$\lim_{N\to\infty} c_N = 1 + \lim_{N\to\infty}\sum_{l=1}^{N} b_l c_{N-l} = 1 + f_{ii}\lim_{N\to\infty} c_N,$$

即
$$c_\infty = 1 + f_{ii}c_\infty \Rightarrow f_{ii} = \frac{c_\infty - 1}{c_\infty} < 1.$$

反之，若 $f_{ii}<1$，则上式表明：$c_\infty<\infty$. 若 $f_{ii}=1$，则上式表明：$c_\infty=\infty$.

引理 7.2.7 设 $i\to j$，则

$$\text{“}f_{ii} = (<)1\text{”} \Longleftrightarrow \text{“}\sum_{n=1}^{\infty} p_{ij}^{(n)} = (<)\infty.\text{”}$$

证 若 $f_{jj}<1$，则按引理 7.2.6，有$\sum\limits_{n=1}^{\infty} p_{jj}^{(n)}<\infty$. 按引理 7.2.3 (i)，得

$$\sum_{n=1}^{\infty} p_{ij}^{(n)} = f_{ij}\cdot\left(1 + \sum_{n=1}^{\infty} p_{jj}^{(n)}\right) < \infty.$$

反之，假定$\sum\limits_{n=1}^{\infty} p_{ij}^{(n)}<\infty$，则按引理 7.2.4，有 $f_{ij}>0$. 于是，按前面的推导，有

$$\infty > f_{ij}^{-1}\cdot\sum_{n=1}^{\infty} p_{ij}^{(n)} = 1 + \sum_{n=1}^{\infty} p_{jj}^{(n)}.$$

按引理 7.2.6，$f_{jj}<1$.

等号情形的等价关系之证完全类似于上面的分析和推理.

引理 7.2.8 若 $i\longleftrightarrow j$，则

$$\text{“}f_{ii} = (<)1\text{”} \Longleftrightarrow \text{“}f_{jj} = (<)1\text{”}.$$

证 按引理 7.2.6，有

$$f_{jj} < 1 \Rightarrow \sum_{n=0}^{\infty} p_{jj}^{(n+l)} < \infty \ (\forall\, l \geqslant 1).$$

按 Ch.-K. 方程，有

$$\sum_{n=0}^{N} p_{jj}^{(n+l)} = \sum_{n=0}^{N}\sum_{u\in E} p_{ju}^{(n)}\cdot p_{uj}^{(l)} \geqslant \sum_{n=0}^{N} p_{ji}^{(n)} p_{ij}^{(l)} \ (\forall\, N \geqslant 0).$$

这表明：若$\sum_{n=0}^{\infty} p_{ji}^{(n)}=\infty$，则必有 $p_{ij}^{(l)}=0$ ($\forall\ l\geqslant 1$)．于是$i \nrightarrow j$．这与假设$i\rightarrow j$矛盾，故

$$f_{jj}<1 \Rightarrow \sum_{n=0}^{\infty} p_{jj}^{(n)}<\infty.$$

按引理 7.2.7，有 $f_{ii}<1$.

同理可证：$f_{ii}<1 \Rightarrow f_{jj}<1$．综合起来，得 $f_{ii}<1 \Leftrightarrow f_{jj}<1$.

定义 7.2.4 若 $f_{ii}=1$，则称状态i是常返的．若 $f_{ii}<1$，则称状态i是非常返的．若

$$\mu_i=E(\sigma_i/\xi_0=i)<\infty,$$

则称i为正常返状态．若 $f_{ii}=1$，且

$$\mu_i=E(\sigma_i/\xi_0=i)=\infty,$$

则称i为0常返状态.

定理 7.2.1 若状态j是常返的，则系统从状态j出发以概率1无限多次返回状态j；若状态j是非常返的，则系统从任意状态i出发以概率1只能有限次返回状态j；若状态j是常返的，则系统从状态i出发以概率f_{ij}无限多次返回状态j.

证 设

$$A_{ij}(m)=\{\text{系统从状态 } i \text{ 出发至少 } m \text{ 次访问状态 } j\}.$$

则 $A_{ij}(m)=A_{ij}(m)\cap(\sigma_j<\infty)=\bigcup_{k=1}^{\infty} A_{ij}(m)\cap(\sigma_j=k)$.

令 $Q_{ij}(m)=P(A_{ij}(m)/\xi_0=i)$，则$\{Q_{ij}(m), m\geqslant 1, i、j\in E\}$具有下列性质：

$$\text{Q.1}\quad Q_{ij}(m)=\sum_{k=1}^{\infty} f_{ij}^{(k)} Q_{jj}(m-1)=f_{ij}\cdot Q_{jj}(m-1);$$

$$\begin{aligned}(Q_{ij}(m)&=\sum_{k=1}^{\infty} P(A_{ij}(m)\cap(\sigma_j=k)/\xi_0=i)\\&=\sum_{k=1}^{\infty} P(A_{ij}(m)/\xi_0=i,\sigma_j=k)\cdot P(\sigma_j=k/\xi_0=i)\\&=\sum_{k=1}^{\infty} Q_{jj}(m-1)f_{ij}^{(k)});\end{aligned}$$

Q. 2　$Q_{ij}(1)=f_{ij}$. ($\because A_{ij}(1)=\{\xi_0=i\}\cap(\sigma_j<\infty)$);

Q. 3　$Q_{jj}(m)=f_{jj}\cdot Q_{jj}\cdot(m-1)$;

$$Q_{ij}(m)=f_{ij}Q_{jj}(m-1)=f_{ij}\cdot f_{jj}^{m-1}.$$

Q. 4　$Q_{ij}=\lim\limits_{m\to\infty}Q_{ij}(m)$存在. ($\because(Q_{ij}(m),m\geqslant1)$单调不增).

极限 Q_{ij}实际上表示"系统从状态 i 出发无限多次访问状态 j"的概率. 按上述各性质,得

$$Q_{ij}=f_{ij}\cdot\lim_{m\to\infty}f_{jj}^{m}.$$

若 j 常返,则 $f_{jj}=1$,从而,$Q_{ij}=f_{ij}(\forall\ i\in E)$,且 $Q_{jj}=1$.

若 j 非常返,则 $f_{jj}<1$. 从而,$Q_{ij}=0\ (\forall\ i\in E)$.

定理 7. 2. 2　若状态 i 常返,且 $i\to j$,则 $f_{ji}=1$.

证　设

$$A_i(k)=\{\text{系统在时刻 }k\text{ 之后无限多次访问状态 }i\}.$$

则按 Markov 性及定理 7. 2. 1,有

$$\begin{aligned}P(A_i(\sigma_j)/\xi_0=i,\sigma_j=k)&=P(A_i(k)/\xi_k=j)\\&=P(A_i(0)/\xi_0=j)=f_{ji}.\end{aligned}$$

按定理 7. 2. 1,状态 i 常返表明:

$$\begin{aligned}&P(A_i(k)/\xi_0=i)=1\\&\Rightarrow\{\xi_0=i\}\cap(\sigma_j=k)=\{\xi_0=i\}\cap(\sigma_j=k)\cap A_i(k)\\&\Rightarrow P(\sigma_j=k/\xi_0=i)=P(\sigma_j=k,A_i(k)/\xi_0=i)\\&\qquad=P(\sigma_j=k,A_i(\sigma_j)/\xi_0=i)\\&\qquad=P(A_i(\sigma_j)/\xi_0=i,\sigma_j=k)\cdot P(\sigma_j=k/\xi_0=i)\\&\qquad=f_{ji}\cdot f_{ij}^{(k)}.\end{aligned}$$

即　$$f_{ij}^{(k)}=f_{ji}\cdot f_{ij}^{(k)}(\forall\ k\geqslant1).$$

对此式两边关于 k 求和,得

$$f_{ij}=f_{ji}\cdot f_{ij}.$$

按条件 $i\to j$ 及引理 7. 2. 4,有 $f_{ij}>0$,故 $f_{ji}=1$.

定理 7. 2. 3

(i) 状态 j 常返 $\Longleftrightarrow f_{jj}=1 \Longleftrightarrow \sum_{n=1}^{\infty} p_{jj}^{(n)}=\infty$；

(ii) $i \longleftrightarrow j \Rightarrow i, j$ 具有相同的常返性，即要么同时常返，要么同时非常返.

证 结论(i)是定义 7.2.4 与引理 7.2.6 的推论. 按引理 7.2.8，有 $f_{ii}=1 \Longleftrightarrow f_{jj}=1$. 从而，$i$、$j$ 具有相同的常返性.

注 1 关于互通的常返状态，更重要的性质将在下面讨论. 并将得到结论：当 $i \longleftrightarrow j$ 时，若 i、j 为常返状态，则它们同是正常返或同是零常返. 这涉及到首达时 σ_i 的平均值 μ_i.

定义 7.2.5 设

$$d=\max\{t: p_{ii}^{(n)}=0 \text{ 当 } n \neq kt, k=0,1,2,\cdots, \text{时}\}.$$

若 $1<d<\infty$，则称 d 为状态 i 的周期.

若 $d=1$，则称状态 i 为非周期.

若 $d=+\infty$，则称状态 i 具有无限周期.

若 $d=1$，且 $\mu_i=E(\sigma_i/\xi_0=i)=E(\sigma_i/\xi_0)|_{\xi_0=i}<\infty$，则称状态 i 是遍历的，即遍历状态为非周期正常返状态.

注 2 如果状态 i 具有无限周期，则必有 $p_{ii}^{(n)}=0$. 这意味着：状态 i 是非常返的. 由此可见，任意常返状态都具有有限的周期.

讨论转移概率 $p_{ij}^{(n)}$ 的渐近行为问题时，状态的周期性及首达时的条件期望起着特别重要的作用. 下面就是研究这些量之间的关系. 为此，首先介绍几条引理.

引理 7.2.9 设 d 为正整数集$(n_l, 1 \leqslant l \leqslant s)$的最大公因子，则 $\exists$ 整数 $m_0>0$，使得方程

$$md=\sum_{i=1}^{s} b_i n_i$$

对任意整数 $m \geqslant m_0$ 存在一个非负整数解$(b_i, 1 \leqslant i \leqslant s)$. (这里不考虑唯一性).

证 设

$$A=\{x: x=\sum_{i=1}^{s} a_i n_i, \text{其中 } a_i \text{ 为任意整数(可正，可负)},$$

$1 \leqslant i \leqslant s\}$.

则 A 具有下列性质：

A.1　若 $x \in A$，则 x 能被 d 整除.

A.2　设 d_0 为 A 中的最小正整数，则对 $\forall\ x \in A$，及整数 k，有 $x - kd_0 \in A$. 事实上，$d_0 \in A$ 表明：$d_0 = \sum_{i=1}^{s} a_i^0 n_i$. $x \in A$ 表明：$x = \sum_{i=1}^{s} a_i n_i$. 因此，

$$x - d_0 k = \sum_{i=1}^{s}(a_i - ka_i^0)n_i \in A.$$

A.3.　对 $\forall\ x \in A$，$\exists$ 整数 k，$\ni x = kd_0$，采用反证法证明这一事实. 假如，对 $\forall$ 整数 k，有 $x \neq kd_0$. 令

$$A_1 = \{x - kd_0 ; k \text{ 为任意整数}\} \subset A.$$

并设 d'_0 为 A_1 的最小正整数，则显然有 $d'_0 \geqslant d_0$. 若 $d'_0 = d_0$，则因 $d'_0 = x - k'd_0$ 而有 $x = (k'+1)d_0$. 这与反假设不合. 于是，$d'_0 > d_0$. 然而，$d'_0 - d_0 \in A$，且为正整数. 因此，应有 $d'_0 - d_0 \geqslant d_0$. 同理可证：此式中等号也不成立. 于是，$d'_0 > 2d_0$. 依此类推，便有 $d'_0 > nd_0$ ($\forall$ 整数 $n \geqslant 1$). 按 d_0 和 d'_0 之定义，这是不可能的. 故 $d'_0 = d_0$.

最后的这条性质表明：d_0 为 A 的最大公因子.

现在，设

$$B = \left\{x : x = \sum_{i=1}^{s} b_i n_i,\text{其中 } b_i \text{ 为非负整数}, 1 \leqslant i \leqslant s\right\}.$$

则 B 具有下列性质：

B.1.　$d_0 = N_1 - N_2$，其中 $N_1, N_2 \in B$. 事实上，$d_0 \in A$ 表明：

$$d_0 = \sum_{i=1}^{s} c_i n_i = \sum_{i=1}^{s} c_i^+ n_i - \sum_{i=1}^{s} c_i^- n_i = N_1 - N_2,$$

其中 $N_1 = \sum_{i=1}^{s} c_i^+ n_i$, $N_2 = \sum_{i=1}^{s} c_i^- n_i$, $c_i^+ = c_i \vee 0$, $c_i^- = c_i^+ - c_i$.

B.2　记 $d_1 = \sum_{i=1}^{s} n_i$, $c = \max\{c_i^- ; 1 \leqslant i \leqslant s\}$，对 $\forall$ 正整数 m，有

表示：

$$m = kd_1 + m_1,$$

其中 k 为非负整数，$1 \leqslant m_1 \leqslant d_1 - 1$. 记为 $m = (k, m_1)$. 若 $kd_0 > m_1 c$，则 $md_0 \in B$. 事实上

$$md_0 = kd_0 d_1 + m_1 d_0 = m_1 N_1 + \sum_{i=1}^{s} (kd_0 - m_1 c_i^-) \cdot n_i$$

按条件，$kd_0 > m_1 c \geqslant m_1 c_i^- (1 \leqslant i \leqslant s) \Rightarrow md_0 \in B$.

B.3　选取 $m^* = (k^*, 0)$，其中 $k^* = 1 + \frac{d_1 c}{d_0}$，则

$$k^* d_0 = d_1 c + d_0 > d_1 c.$$

因此，若整数 $m = (k, m_1) \geqslant m^*$，则

$$kd_0 \geqslant k^* d_0 > d_1 c > m_1 c.$$

从而，按集 B 的性质 B.2，有 $md_0 \in B$ ($\forall\ m \geqslant m^*$).

为了完成引理之证，最后余下的问题是构作 m_0. 按 d_0 之定义及 $n_i \in A$，d_0 为 $(n_i, 1 \leqslant i \leqslant s)$ 的公因子. 因此，$d_0 \leqslant d$，且 d 能被 d_0 整除，即 $\exists\ k_0 \geqslant 1$，$\ni$ $d = k_0 d_0$. 现在，选取 $m_0 \geqslant \frac{m^*}{k_0}$，则当 $m \geqslant m_0$ 时，有

$$mk_0 \geqslant k_0 m_0 \geqslant m^*$$

$$\Rightarrow md = mk_0 d_0 \in B\ (\forall\ m \geqslant m_0)\ (\text{按 B.3}).$$

引理 7.2.10　设 d' 为集 $\{n \geqslant 1: f_{ii}^{(n)} > 0\}$ 的最大公因子，则 $d' = d$，其中 d 为状态 i 的周期(此结论表明：集 $\{n \geqslant 1: f_{ii}^{(n)} > 0\}$ 与 $\{n \geqslant 1: p_{ii}^{(n)} > 0\}$ 有相同的周期或最大公因子).

证　按引理 7.2.3(i)，有 $p_{ii}^{(n)} \geqslant f_{ii}^{(n)}$. 因此有

$$\{n \geqslant 1: f_{ii}^{(n)} > 0\} \subset \{n \geqslant 1: p_{ii}^{(n)} > 0\} \Rightarrow d' \geqslant d.$$

于是，结论等价于证明：$d' \leqslant d$. 若 $d' = 1$，则自然有 $d' \leqslant d$. 若 $d' = \infty$，则 $f_{ii}^{(n)} = 0$ ($\forall\ n \geqslant 1$). 按引理 7.2.3(i)，必有 $p_{ii}^{(n)} = 0$ ($\forall\ n \geqslant 1$). 从而，$d = \infty$. 于是，亦有 $d' \leqslant d$. 最后，考虑 $1 < d' < \infty$ 的情形，按 d' 之定义，有 $f_{ii}^{(kd'+r)} = 0$ ($\forall\ k \geqslant 0$; $1 \leqslant r \leqslant d' - 1$). 利用这一事实及引

理 7.2.3(i)，便可推出：对于 $1\leqslant r\leqslant d'-1$，有

$$p_{ii}^{(r)}=\sum_{l=1}^{r}f_{ii}^{(l)}p_{ii}^{(r-l)}=0,$$

$$p_{ii}^{(d'+r)}=\sum_{l=0}^{r}f_{ii}^{(d'+l)}\cdot p_{ii}^{(r-l)}=0.$$

$$p_{ii}^{(kd'+r)}=0\ (\forall\ k\geqslant 0),$$

即 $p_{ii}^{(n)}=0$ 当 $n\neq kd'$，$k=0,1,2,\cdots$，时.

故 d' 亦为 $\{n\geqslant 1:p_{ii}^{(n)}>0\}$ 的一个公因子. 从而，$d'\leqslant d$.

定理 7.2.4 设 d 为状态 i 的周期(含 $d=1$)，则∃ 正整数 k^*，⇁ $p_{ii}^{(kd)}>0\ (\forall\ k\geqslant k^*)$.

证 设

$$A=\{n_l:l\geqslant 1\}=\{n\geqslant 1:p_{ii}^{(n)}>0\},\text{其中 } n_l\uparrow.$$

按假设，d 为 A 的最大公因子. 现假定，t_l 为 $\{n_1,n_2,\cdots,n_l\}$ 的最大公因子；则 $t_l\downarrow$，且 $t_l\geqslant d\ (\forall\ l\geqslant 1)$. 显然，$t_l$ 能被 d 和 t_{l+m} 整除($\forall\ m\geqslant 0,l\geqslant 1$)，注意，$n_1$ 为一有限数. 因此，n_1 只能具有有限个质因子，且 n_1 能被所有的 t_l 整除. 若在集 $\{t_l,l\geqslant 1\}$ 中有无限多个元素不同，则 n_1 就会有无限多个不同的质因子. 这不可能. 因此，∃ $l_0\geqslant 1$，⇁ $t_l=t_{l_0+m}=d\ (\forall\ l\geqslant l_0,m\geqslant 0)$. 这表明：$d$ 为 $\{n_1,n_2,\cdots,n_{l_0}\}$ 的最大公因子. 按引理 7.2.9，∃ 整数 $k^*>0$，⇁当 $k\geqslant k^*$ 时，有表示：

$$kd=\sum_{i=1}^{l_0}b_in_i\ (b_i\text{ 为非负整数},1\leqslant i\leqslant l_0).$$

按 Ch.-K. 方程，有

$$p_{ii}^{(kd)}\geqslant\prod_{l=1}^{l_0}p_{ii}^{(b_ln_l)}\geqslant\prod_{l=1}^{l_0}(p_{ii}^{(n_l)})^{b_l}>0\ (\forall\ k\geqslant k^*).$$

定理 7.2.5

(i) 若 j 为正常返状态，且周期为 d(含 $d=1$)，则

$$\lim_{n\to\infty}p_{jj}^{(nd)}=\frac{d}{\mu_j},\text{且}\overline{\lim_{n\to\infty}}p_{jj}^{(n)}=\frac{d}{\mu_j}.$$

(ii)若 j 为 0 常返状态，则 $\lim\limits_{n\to\infty}p_{jj}^{(n)}=0$.

证 为简化表写，省去下标"j"，如

$$f^{(n)}=f_{jj}^{(n)},p^{(n)}=p_{jj}^{(n)},\mu=\mu_j,f=f_{jj}.$$

定义符号：$r_n=\sum_{k=n+1}^{\infty}f^{(k)}(\forall\ n\geqslant 0)$.

首先讨论：j 为正常返，周期为 d（含 $d=1$）的情形. 数列 $r=(r_n,n\geqslant 0)$ 具有下列性质：

R. 1 $r_0=\sum_{k=1}^{\infty}f^{(k)}=f=1$.（由 j 的常返性）.

R. 2 $\sum_{l=0}^{n}r_l p^{(n-l)}=1\ (\forall\ n\geqslant 0)$. 事实上，按引理 7. 2. 3(i)，有

$$p^{(n)}=\sum_{l=1}^{n}f^{(l)}p^{(n-l)}=\sum_{l=1}^{n}(r_{l-1}-r_l)\cdot p^{(n-l)}$$

$$\Rightarrow\sum_{l=0}^{n}r_l p^{(n-l)}=\sum_{l=0}^{n-1}r_l p^{(n-1-l)}=\cdots=r_0\cdot p^{(0)}=1.$$

R. 3 $$\sum_{n=0}^{\infty}r_n=\sum_{n=1}^{\infty}nr_{n-1}-\sum_{n=1}^{\infty}nr_n=\sum_{n=1}^{\infty}n(r_{n-1}-r_n)$$

$$=\sum_{n=1}^{\infty}nf^{(n)}=\mu.$$

现在，令

$$\lambda=\overline{\lim_{n\to\infty}}p^{(nd)}=\overline{\lim_{n\to\infty}}p^{(n)},$$

则 λ 具有下列几种特别表示：

Λ. 1 ∃ 子列$(n_l,l\geqslant 1)$ $(n_l\uparrow)$，使得

$$\lambda=\lim_{l\to\infty}p^{(n_l d-s)}(\forall\ s\in\{n\geqslant 1;f^{(n)}>0\}).$$

事实上，按 λ 之定义，∃ 子列$\{n_l,l\geqslant 1\}$ $(n_l\uparrow)$，使得

$$\lambda=\lim_{l\to\infty}p^{(n_l d)}=\overline{\lim_{n\to\infty}}p^{(n_l d)}.$$

于是，按引理 7. 2. 3(i)和 7. 2. 5，有

$$\lambda=\underline{\lim_{l\to\infty}}(f^{(s)}\cdot p^{(n_l d-s)}+\sum_{\substack{\nu=1\\ \nu\neq s}}^{n_l d}f^{(\nu)}p^{(n_l d-\nu)})$$

$$\leqslant f^{(s)} \varliminf_{i\to\infty} p^{(n_i d-s)} + \sum_{\substack{\nu=1\\ \nu\neq s}}^{\infty} f^{(\nu)} \cdot \varlimsup_{m\to\infty} p^{(m)}$$

$$= f^{(s)} \varliminf_{i\to\infty} p^{(n_i d-s)} + (1 - f^{(s)})\lambda$$

$$\Rightarrow \lambda \leqslant \varliminf_{i\to\infty} p^{(n_i d-s)} \leqslant \lambda \ (\forall\ f^{(s)} > 0).$$

Λ. 2 $\exists\ k_0 \geqslant 1$, 使 $\lim\limits_{i\to\infty} p^{(n_i d-kd)} = \lambda \ (\forall\ k \geqslant k_0)$.

事实上，按引理 7.2.10，d 为 $\{n\geqslant 1: f^{(n)}>0\}$ 的最大公因子. 设 $(m_l, l\geqslant 1) = (n\geqslant 1: f^{(n)}>0)$ $(m_l \uparrow)$. 类似于定理 7.2.4 之证，可以找到一个整数 $l_0 \geqslant 1$ 及整数 k_0，使得，对$\forall\ k\geqslant k_0$，有

$$kd = \sum_{l=1}^{l_0} c_l m_l, \text{其中 } c_l \text{ 为非负整数}, 1 \leqslant l \leqslant l_0.$$

不妨设每个 $c_l>0$（如果$=0$，就从中删去）. 若让

$$n'_k d = n_k d - m_1$$

（此式的合理性在于 $m_1 \in \{n\geqslant 1: f^{(n)}>0\}$），则按 Λ. 1，有

$$\lambda = \lim_{k\to\infty} p^{(n'_k d)}.$$

同理，让 $n''_k d = n'_k d - m_1$，有 $\lambda = \lim\limits_{k\to\infty} p^{(n''_k d)}$. 依此类推，便可得：$\lambda = \lim\limits_{k\to\infty} p^{(n_k d - c_1 m_1)}$. 类似地推理，可得

$$\lambda = \lim_{k\to\infty} p^{(n_k d - c_1 m_1 - c_2 m_2)}.$$

于是，最终可得

$$\lim_{i\to\infty} p^{(n_i d-kd)} = \lim_{i\to\infty} p^{(n_i d - \sum_{l=1}^{l_0} c_l m_l)} = \lambda \ (\forall\ k \geqslant k_0).$$

Λ. 3 $\lambda = \left(\sum\limits_{\nu=0}^{\infty} r_{\nu d}\right)^{-1}$.

事实上，考虑到 $p^{(kd+m)}=0$ $(\forall\ k\geqslant 0, 1\leqslant m\leqslant d-1)$，按 R. 2，有

$$\sum_{\nu=0}^{n_i-k} r_{\nu d} p^{((n_i-k-\nu)d)} = 1 \ (\forall\ i \geqslant 1, k \geqslant 0).$$

在此式中让 $i\to\infty$ 求极限，则按引理 7.2.5 及 Λ. 2，得

$$\sum_{\nu=0}^{\infty} r_{\nu d} \cdot \lambda = 1.$$

A.4 $\lambda=\frac{d}{\mu}$.

事实上，令 $q=\sum_{\nu=0}^{\infty}r_{\nu d}$. 考虑到：$f^{(\nu)}=0\ (\nu\neq kd, k\geqslant 0)$. 则

$$r_{\nu d}=\frac{1}{d}\sum_{k=\nu d}^{(\nu+1)d-1}r_k$$

$$\Rightarrow q=\frac{1}{d}\sum_{\nu=0}^{\infty}\sum_{k=\nu d}^{(\nu+1)d-1}r_k=\frac{1}{d}\sum_{k=0}^{\infty}r_k=\frac{\mu}{d}$$

$$\Rightarrow \lambda=\frac{1}{q}=\frac{d}{\mu}.$$

A.5 令 $\lambda^*=\varliminf_{n\to\infty}p^{(nd)}$，则 $\lambda^*=\frac{d}{\mu}=\lambda$.

事实上，按定理 7.2.4，$\exists\ n_0\geqslant 1$，$\ni p^{(nd)}>0\ (\forall\ n\geqslant n_0)$.

因此，$\{p^{(nd)}, n\geqslant 1\}$ 中不会有子列的元素全为 0. 于是，可选子列 $(n_l, l\geqslant 1)$，使得

$$\lambda^*=\varliminf_{n\to\infty}p^{(nd)}=\lim_{k\to\infty}p^{(n_k d)}=\varlimsup_{k\to\infty}p^{(n_k d)}.$$

从对 λ 的结构的分析中可以看出：$\lambda^*=\lambda$ 等价于：

$$\lambda^*=\lim_{i\to\infty}p^{(n_i d-s)}\ (\forall\ f^{(s)}>0).$$

按引理 7.2.10，上式中的 s 有表示 $s=md$. 于是 $\lambda^*=\lambda$ 等价于：

$$\lambda^*=\lim_{i\to\infty}p^{(n_i d-md)}\ (\forall\ f^{(md)}>0).$$

为证此式，需要下列估计：

$$\lambda^*=\varliminf_{i\to\infty}\left(f^{(md)}\cdot p^{(n_i d-md)}+\sum_{\substack{\nu=1\\ \nu\neq m}}^{n_i}f^{(\nu d)}p^{(n_i d-\nu d)}\right)\ (\text{按引理 } 7.2.3(\mathrm{i}))$$

$$\geqslant f^{(md)}\varliminf_{i\to\infty}p^{(n_i d-md)}+\varliminf_{i\to\infty}\sum_{\substack{\nu=1\\ \nu\neq m}}^{n_i}f^{(\nu d)}p^{(n_i d-\nu d)}$$

$$\geqslant \text{同第一项}+\sum_{\substack{\nu=1\\ \nu\neq m}}^{\infty}f^{(\nu d)}\varliminf_{i\to\infty}p^{(n_i d-\nu d)}\ (\text{按引理 } 7.2.5)$$

$$\geqslant \text{同第一项}+(1-f^{(md)})\lambda^*.$$

$$\Rightarrow \lambda^* \geqslant \overline{\lim_{i\to\infty}} p^{(n_i d - md)} \geqslant \underline{\lim_{i\to\infty}} p^{(n_i d - md)} \geqslant \lambda^*. \ (\forall\ f^{(md)} > 0).$$

此后的证明:完全与 λ 的相同.故 $\lambda^* = \dfrac{d}{\mu} = \lambda$.

当 $d=1$ 时,考虑到 $\{n\geqslant 1: f^{(n)}>0\}$ 与 $\{n\geqslant 1: p^{(n)}>0\}$ 有相同的最大公因子 1.便可知:上述分析完全适用于这种情形.

综合起来得结论(i).

最后考证结论(ii).设 j 为 0 常返状态,则

$$\sum_{k=0}^{nd} r_k = d \cdot \sum_{\nu=0}^{n} r_{\nu d}, \text{且 } \mu = \infty$$

$$\Rightarrow \lim_{n\to\infty} \sum_{\nu=0}^{n} r_{\nu d} = \infty$$

令 $\lambda = \overline{\lim_{n\to\infty}} p^{(nd)}$,则跟前面的讨论一样.

存在子列 $\{n_l: l\geqslant 1\} \subset \{n_l \geqslant 1: p^{(n)}>0\}$ 及 $k_0 \geqslant 1$,使得

$$\lambda = \lim_{i\to\infty} p^{(n_i d)} = \lim_{i\to\infty} p^{(n_i d - kd)} (\forall\ k \geqslant k_0).$$

令 $r_n^* = \sum_{\nu=0}^{n} r_{\nu d}$,则按 R. 2,及引理 7.2.5,可推出:

$$0 = \lim_{i\to\infty} \frac{1}{r_{n_i}^*} = \lim_{i\to\infty} \frac{1}{r_{n_i}^*} \sum_{\nu=0}^{n_i} r_{\nu d} p^{(n_i d - \nu d)}$$

$$= \lim_{i\to\infty} \left(\frac{1}{r_{n_i}^*} \sum_{\nu=0}^{k_0} r_{\nu d} p^{(n_i d - \nu d)} \right) + \lim_{i\to\infty} \left(\frac{1}{r_{n_i}^*} \sum_{\nu=k_0+1}^{n_i} r_{\nu d} p^{(n_i d - \nu d)} \right)$$

$$= \lim_{i\to\infty} p^{(n_i d)}$$

$$\Rightarrow \lambda = \overline{\lim_{n\to\infty}} p^{(nd)} = 0.$$

状态判据:

(Ⅰ)状态 i 正常返 $\Longleftrightarrow \mu_i = E(\sigma_i/\xi_0 = i) < \infty$

$$\Longleftrightarrow \sum_{n=1}^{\infty} p_{ii}^{(n)} = \infty, \text{且} \underline{\lim_{n\to\infty}} p_{ii}^{(n)} > 0.$$

(Ⅱ)状态 i 零常返 $\Longleftrightarrow f_{ii} = 1, \mu_i = \infty$

$$\Longleftrightarrow \sum_{n=1}^{\infty} p_{ii}^{(n)} = \infty, \text{且} \overline{\lim_{n\to\infty}} p_{ii}^{(n)} = 0.$$

(Ⅲ)状态 i 非常返 $\Leftrightarrow f_{ii}<1 \Leftrightarrow \sum_{n=1}^{\infty} p_{ii}^{(n)} < \infty$.

定理 7.2.6 设 $i \longleftrightarrow j$，则

(i) i 和 j 同时常返或同时非常返；

(ii) 当 i 和 j 同时常返时，要么同时正常返，要么同时零常返；

(iii) i 和 j 具有相同的周期性，即 $d_i=d_j$，其中 d_k 表示状态 k 的周期(含 $d_i=1=d_j$).

证 首先由条件 $i \longleftrightarrow j$ 可得如下基本关系(7.2-1)和(7.2-2). 注意，$i \longleftrightarrow j$ 表明：$\exists\ l \geqslant 1$ 和 $n \geqslant 1$，使得

$$p_{ij}^{(l)} = \alpha > 0,\ p_{ji}^{(n)} = \beta > 0.$$

按 Ch.-K. 方程便可推出：

$$p_{ii}^{(l+m+n)} \geqslant p_{ij}^{(l)} p_{jj}^{(m)} p_{ji}^{(n)} = \alpha\beta p_{jj}^{(m)} (\forall\ m \in N); \quad (7.2\text{-}1)$$

$$p_{jj}^{(l+m+n)} \geqslant p_{ji}^{(n)} \cdot p_{ii}^{(m)} \cdot p_{ij}^{(l)} = \alpha\beta p_{ii}^{(m)}.\ (\forall\ m \in N). \quad (7.2\text{-}2)$$

结论(i)即为定理 7.2.3.

为证结论(ii)，先设 i 为正常返，则

$$\overline{\lim_{m\to\infty}} p_{jj}^{(m)} = \overline{\lim_{m\to\infty}} p_{jj}^{(l+m+n)} \geqslant \alpha\beta\, \overline{\lim_{m\to\infty}} p_{ii}^{(m)} > 0.$$

按引理 7.2.8 和判据(Ⅰ)，j 为正常返. 现设 i 为零常返，则

$$0 = \overline{\lim_{m\to\infty}} p_{ii}^{(m)} \geqslant \alpha\beta\, \overline{\lim_{m\to\infty}} p_{jj}^{(m)}.$$

从而，j 为零常返. 按对称性，得结论(ii).

为证结论(iii)，设 $m \in \{n \geqslant 1 : p_{jj}^{(n)} > 0\}$，则按(7.2-1)式，有

$$l + m + n \in \{n \geqslant 1 : p_{ii}^{(n)} > 0\}.$$

于是，存在整数 $k_1 \geqslant 1$，使得 $l+m+n=k_1 d_i$. 在(7.2-1)式中，让 $m=0$，则得 $l+n \in \{n \geqslant 1 : p_{ii}^{(n)} > 0\}$. 于是，$\exists$ 整数 $k_2 \geqslant 1$，$\ni l+n = k_2 d_i$. 综合起来，得：对 $\forall\ m \in \{n \geqslant 1 : p_{jj}^{(n)} > 0\}$，$\exists\ k_1, k_2 \geqslant 1$，使得

$$m = (k_1 - k_2) d_i.$$

故 d_i 为 $\{n\geqslant 1: p_{jj}^{(n)}>0\}$ 的公因子. 从而，得 $d_j\geqslant d_i$. 按对称性，得 $d_i \geqslant d_j$. 故 $d_i=d_j$.

推论 7.2.1 设 j 为常返状态. 周期为 d(含 $d=1$)，则

$$\overline{\lim_{n\to\infty}} p_{ij}^{(n)} \leqslant f_{ij}\frac{d}{\mu_j} \ (\forall\, i\in E).$$

若 $d=1$，则 $\lim\limits_{n\to\infty} p_{ij}^{(n)}=\dfrac{f_{ij}}{\mu_j}$ $(\forall\, i\in E)$.

证 假如 $f_{ij}=0$，则按引理 7.2.3(i)，有

$$p_{ij}^{(n)}=0 \ (\forall\, n\geqslant 1).$$

从而，推论成立.

现设 $f_{ij}>0$，则按引理 7.2.3(i)，引理 7.2.5 及定理 7.2.5，有

$$\begin{aligned}\overline{\lim_{n\to\infty}} p_{ij}^{(n)} &= \overline{\lim_{n\to\infty}}\sum_{k=1}^{n} f_{ij}^{(k)} p_{jj}^{(n-k)} \\ &\leqslant \lim_{n\to\infty}\sum_{k=1}^{n} f_{ij}^{(k)} \sup\{p_{jj}^{(\nu)}, \nu\geqslant n-k\} \\ &= f_{ij}\,\overline{\lim_{n\to\infty}} p_{jj}^{(n)} = f_{ij}\cdot\frac{d}{\mu_j}.\end{aligned}$$

若 $d=1$，则按定理 7.2.5，极限 $\lim\limits_{n\to\infty} p_{jj}^{(n)}$ 存在. 于是，按引理 7.2.5，有

$$\begin{aligned}\lim_{n\to\infty} p_{ij}^{(n)} &= \lim_{n\to\infty}\sum_{k=1}^{n} f_{ij}^{(k)} p_{jj}^{(n-k)} \\ &= f_{ij}\lim_{n\to\infty} p_{jj}^{(n)} = f_{ij}\frac{1}{\mu_j}.\end{aligned}$$

§7.3 闭集和状态空间的分解

上节介绍状态空间 E 中元素的性质. 本节将在此基础上讨论状态空间 E 的分解. 所有采用的基本符号均跟上节一致.

定义 7.3.1 (闭集) 设 C 为状态空间 E 中的一个子集. 若对 $\forall\, i\in C$，及 $j\notin C$，有 $p_{ij}=0$，则称 C 为闭集.

注1 C 为闭集. 但 $E\backslash C$ 不必为闭集. 设 $E=\{0,1,2\}$ 的转移图为图 7.3-1,转移阵为

$$P=\begin{array}{c|c} \begin{array}{ccc} 0 & 1 & 2 \end{array} & \\ \hline \begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix} & \begin{array}{c} 0 \\ 1 \\ 2 \end{array} \end{array}$$

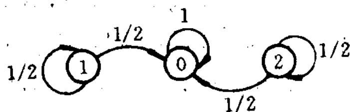


图 7.3-1 转移图

令 $C_1=\{0\},C_2=\{0,1\},C_3=\{0,2\}$. 则 C_i 均为闭集,但 $E\backslash C_i$ 均非闭集. E 本身是最大的闭集.

引理 7.3.1 C 为闭集的充要条件是下列两条件之一成立:

(i) $i\in C,j\notin C\Rightarrow i\nrightarrow j$;

(ii) $\sum\limits_{j\in C}p_{ij}^{(n)}=1\ (\forall\ i\in C,n\in\mathcal{N})$.

证 C 为闭集表明:$i\in C,j\notin C\Rightarrow p_{ij}=0$. 按 Ch.-K. 方程,可得

$$p_{ij}^{(2)}=\sum_{s\in E}p_{is}p_{sj}=\sum_{s\in C}p_{is}p_{sj}+\sum_{s\notin C}p_{is}p_{sj}=0.$$

假如 $p_{is}^{(m)}=0(\forall\ s\notin C)$,则按 Ch.-K. 方程,有

$$p_{ij}^{(m+1)}=\sum_{s\in C}p_{is}^{(m)}p_{sj}+\sum_{s\notin C}p_{is}^{(m)}p_{sj}=0$$

$$\Rightarrow p_{ij}^{(n)}=0\ (\forall\ n\in \mathrm{N})\Rightarrow i\nrightarrow j,\text{即条件(i) 成立.}$$

由条件(i)立即可得条件(ii).

反之,若(i)成立,则 $p_{ij}=0(i\in C,j\notin C)$. 从而,$C$ 为闭集. 显然,条件(i)和(ii)等价.

注2 按引理 7.1.1,若 C 为闭集,则 $\mathrm{P}_C=[p_{ij}]_{i,j\in C}$ 为转移阵,且 $\mathrm{P}_C^{(n)}=\mathrm{P}_C^n$. $(\forall\ n\in\mathrm{N})$. 这个事实表明:当 C 为 E 中之闭子集时,将状态空间 E 限制到 C 后,ξ 必为以 C 为状态空间的齐次,

Markov 链. 因此,从 E 中分解出闭集就是一个很基本的问题.

定理 7.3.1 状态空间 E 中所有常返状态构成一个闭集.

证 设 C 为 E 中常返状态全体. 若 $C=E$,则 C 自然为闭集. 现假定 $C\neq E$,且不为闭集. 则按定义 7.3.1,∃ $i\in C, j\notin C, \ni p_{ij}\neq 0$. 这表明: $i\to j$. 由于 j 为非常返状态而有 $f_{jj}<1$. 按引理 7.2.7, 有 $\sum_n p_{ij}^{(n)}<\infty$. 然而,按 Ch.-K. 方程,有

$$\sum_{n=1}^{N} p_{ij}^{(n)} = \sum_{n=1}^{N}\sum_{s\in E} p_{is}^{(n-1)} p_{sj} \geqslant p_{ij}\cdot \sum_{n=1}^{N} p_{ii}^{(n-1)}.$$

按定理 7.2.3 及 $p_{ij}>0$,上式右边当 $N\to\infty$ 时趋向于∞. 因此, $\sum_n p_{ij}^{(n)}=\infty$. 这就导致矛盾. 故 C 为闭集.

定义 7.3.2 如果齐次 Markov 链 ξ 的状态空间 E 不含真闭集,则称 ξ 为不可约 Markov 链.

引理 7.3.2 若 E 是不可约的,则 E 要么不含常返状态,要么不含非常返状态.

证 E 不可约表明:它也为最小闭集. 按定理 7.3.1, E 中常返状态全体 C 构成一个闭集. 因此,若 E 含常返状态,则 $E=C$.

到目前为止,已经知道:在 E 中去掉非本质状态后剩下的构成一个闭集. 而在此闭集中常返状态全体构成一个闭集. 从引理 7.3.2 中看出:不可约链是基本链,特别是常返不可约链. 这就是为什么在齐次 Markov 链理论中不可约链是人们研究的基本对象,按引理 7.3.1,不同闭集之间的元素是不能互通的. 这个事实表明:按互通关系定义的状态类必为闭集. 这启示我们: E 中常返状态全体可以按"互通关系"分解.

定义 7.3.2 设 C^* 为 E 中常返状态全体,则存在不可约闭集列 $C=(C_l, l\geqslant 1), \ni C^*=\bigcup_{l=1}^{\infty} C_l$(空集为闭集),其中 C 满足下列要求:

(i) C_l 中的状态互通($\forall\ l\geqslant 1$);

(ii) $C_l\cap C_k=\varnothing\ (l\neq k)$.

而且满足条件(i)和(ii)的分解是唯一的.

证 任取 $i_1 \in C^*$. 定义 $C_1 = \{j \in C^*: j \longleftrightarrow i_1\}$. 任取 $i_2 \in C^* \setminus C_1$,定义

$$C_2 = \{j \in C^* \setminus C_1 : j \longleftrightarrow i_2\}.$$

假定按此格式已作出 $C_1, C_2, \cdots, C_n$,令 $C_n^* = C^* \setminus \bigcup_{k=1}^{n} C_k$,并任取 $i_{n+1} \in C_n^*$,定义

$$C_{n+1} = \{j \in C_n^* : j \longleftrightarrow i_{n+1}\}.$$

于是,从 C^* 中就可构作出集列$(C_l, l \geqslant 1)$. 具有下列特性:

C.1 C_l 中的状态互通.(按引理 7.2.2);

C.2 C_l 是不可约的. 事实上,若不然,则按引理 7.3.1,存在闭集 C'_l, $j \in C_l \setminus C'_l$, $\ni i \not\to j$ $(\forall\ i \in C'_l)$. 这与 C.1 矛盾;

C.3 $C_l \cap C_k = \varnothing (l \neq k)$. 事实上,若 $C_l \cap C_k \neq \varnothing$,则 $i \in C_l \cap C_k$ 意味着:$i_l \longleftrightarrow i \longleftrightarrow i_k$. 按引理 7.2.2,$i_l \longleftrightarrow i_k$. 于是,$i_l \in C_k$. 这与定义矛盾;

C.4 $C^* = \bigcup_{l=1}^{\infty} C_l$. 事实上,$E$ 为至多可数集. 因此,集列 C 至多可数个. 若 $i^* \in C^*$,则从 i^* 出发可以构作一个互通类 C_{*}. 然而,按集列 C 的构作法,必有 $C_{*} \in C$. 故 $\bigcup_l C_l \supset C^*$. 从而 $C^* = \bigcup_l C_l$.

满足条件(i)和(ii)的分解之唯一性是显然的.

推论 7.3.1 (i) $E = I + \bigcup_l C_l$,其中 I 为 E 中非常返状态全体. 而集列$(C_l, l \geqslant 1)$按定理 7.3.2 中定义;

(ii)若 E 不含常返状态,则

$$E \text{ 不可约} \Longleftrightarrow E \text{ 为互通类};$$

(iii)若 E 为常返不可约状态空间,则 E 的所有元素具有相同的有限周期性.

注 3 此推论表明:具有周期 d(含 $d=1$)的常返不可约 Markov 链是常返不可约 Markov 链类型中的最基本形式. 因此之故,这种类型的 Markov 链成为人们的重要研讨对象之一.

引理 7.3.3 设 $i, j \in E$, j 常返且具有周期 d. 给定 $r \in \mathrm{N}$. 若

$\exists\ k_0 \geqslant 1$，使得 $p_{ij}^{(k_0 d+r)}>0$，则$\exists\ m_0 \geqslant 1$，$\ni p_{ij}^{(md+r)}>0\ (\forall\ m \geqslant m_0)$.

证 事实上，按定理 7.2.4，存在 $m^* \geqslant 1$，使得

$$p_{jj}^{(nd)} > 0\ (\forall\ n \geqslant m^*).$$

按 Ch.-K. 方程，有

$$p_{ij}^{(md+r)} = \sum_{s \in E} p_{is}^{(k_0 d+r)} \cdot p_{sj}^{((m-k_0)d)} \geqslant p_{ij}^{(k_0 d+r)} p_{jj}^{((m-k_0)d)}.$$

现在，取 $m_0 = m^* + k_0$，则

$$m \geqslant m_0 \Longleftrightarrow m - k_0 \geqslant m^*.$$

于是，$p_{jj}^{((m-k_0)d)}>0$. 从而，$p_{ij}^{(md+r)}>0\ (\forall\ m \geqslant m_0)$.

引理 7.3.4 设 K 为 E 中具有周期 d 的互通类. 任取 $i \in K$. 令

$$K_r(i) = \{j \in K : \exists\ k_r \geqslant 1, \ni p_{ij}^{(k_r d+r)} > 0\}\ (0 \leqslant r \leqslant d-1).$$

$$K_m(i) = K_r(i)，其中\ m = kd + r = (k,r) = r(\mathrm{mod} d).$$

则(i) $i \in K_0(i)$，且 $K_0(j) = K_0(i)\ (\forall\ j \in K_0(i))$；

(ii) $K_{r_1}(i) \cap K_{r_2}(i) = \varnothing\ (0 \leqslant r_1、r_2 \leqslant d-1$，且 $r_1 \neq r_2)$；

(iii) $K_{r_1(\mathrm{mod} d)}(i) \cap K_{r_2(\mathrm{mod} d)}(i) = \varnothing\ (0 \leqslant r_1、r_2 \leqslant d-1$，且 $r_1 \neq r_2)$；

(iv) $K_m(i) = \{j \in K : \exists\ k_m \geqslant 1, \ni p_{ij}^{(k_m d+m)}>0\}\ (\forall\ m \in \mathrm{N})$.

证 性质(iv)可以从引理 7.3.3 中得出. 性质(iii)可由引理 7.3.3 及性质(ii)推出. 因此，仅需证明(i)和(ii).

先证(i). 设 $j \in k_0(i)$，则$\exists\ K_0 \geqslant 1$，$\ni p_{ij}^{(k_0 d)}>0$. 由 i 常返立即可得 $i \in K_0(\mathrm{i})$. 若 $u \in K_0(j)$，则$\exists\ k^* \geqslant 1$，$\ni p_{ju}^{(k^* d)}>0$. 于是，按 Ch.-K. 方程，有

$$p_{iu}^{((k_0+k^*)d)} \geqslant p_{ij}^{(k_0 d)} \cdot p_{ju}^{(k^* d)} > 0.$$

故 $\qquad u \in K_0(i)$，即 $K_0(j) \subset K_0(i)$.

按对称性，若 $i \in K_0(j)$，则 $K_0(i) \subset K_0(j)$. 从而，$K_0(i) = K_0(j)$. 为证 $i \in K_0(j)$，即$\exists\ l_0 \geqslant 1$，$\ni p_{ji}^{(l_0 d)}>0$. 注意，按假设，$i \longleftrightarrow j$. 因此，$\exists\ n_0 \geqslant 1$，$\ni p_{ji}^{(n_0)}>0$. 从而

$$p_{ii}^{(k_0d+n_0)} \geqslant p_{ij}^{(k_1d)} \cdot p_{ji}^{(n_0)} > 0 \ (\because i, j \in K_0(i)).$$

又 i 具有周期 d，上式表明：k_0d+n_0 能被 d 整除，即∃ $l_0 \geqslant 1$，∋ $n_0 = l_0 d$. 于是，$p_{ji}^{(l_0d)} > 0$. 故 $i \in K_0(j)$.

最后证(ii). 假如 $K_{r_1}(i) \cap K_{r_2}(i) \neq \varnothing$. 让 $j \in K_{r_1}(i) \cap K_{r_2}(i)$，则∃ $k_1 \geqslant 1, k_2 \geqslant 1$，∋ $p_{ij}^{(k_1d+r_1)} > 0, p_{ij}^{(k_2d+r_2)} > 0$. 注意，$K$ 为互通类. 因此，∃ $m \geqslant 1$ ∋ $p_{ji}^{(m)} > 0$.

于是，
$$p_{ii}^{(k_1d+r_1+m)} \geqslant p_{ij}^{(k_1d+r_1)} \cdot p_{ji}^{(m)} > 0.$$

因 i 具有周期 d. 因此，r_1+m 能被 d 整除. 然而，$|r_2-r_1|<d$，故只能有 $r_1=r_2$. 这表明：若 $r_1 \neq r_2$，则

$$K_{r_1}(i) \cap K_{r_2}(i) = \varnothing.$$

定理 7.3.3 设 K 为 E 中具有周期 d 的互通类. 则 K 可分解成 d 个互不相联的子类：

$$K_{0(\mathrm{mod}\,d)}, K_{1(\mathrm{mod}\,d)}, \cdots, K_{(d-1)(\mathrm{mod}\,d)}.$$

使得从 K_m 出发仅一步就能运动到 K_{m+1} ($\forall\, m \in \mathrm{N}$). 此外，对∀ $j_u \in K_{r_u}$ $(u=1,2)$，∃ 整数 $N=N(j_1,j_2)$，∋ $p_{j_1j_2}^{(nd+r_2-r_1)} > 0$ ($\forall\, n \geqslant N$).

证 取 $K_m = K_m(i)$，其中 $K_m(i)$ 按引理 7.3.4 定义. 按此引理，$\{K_{r(\mathrm{mod}\,d)}, 0 \leqslant r < d-1\}$ 互不相联. 易证：$K = \bigcup_{r=0}^{d-1} K_r$. 事实上，对 $\forall\, j \in K$，因 $j \longleftrightarrow i$，而∃ $m \geqslant 1$，∋ $p_{ij}^{(m)} > 0$，故 $j \in K_m$. 又 m 有表示：$m=kd+r$. $(0 \leqslant r \leqslant d-1)$，于是，$j \in K_r \subset \bigcup_{r=0}^{d-1} K_r$，即 $K \subset \bigcup_{r=0}^{d-1} K_r \subset K$.

现在证明：$\sum_{u \in K_r} p_{su} = 1$ ($\forall\, s \in K_{r-1}$). 事实上，按引理 7.3.3，对 $\forall\, u \in K_r$，∃ $k_0 \geqslant 1$，∋ $p_{iu}^{(kd+r)} > 0$ ($\forall\, k \geqslant k_0$). 按 K_r 之定义，当 $u \notin K_r$ 时，$p_{iu}^{(kd+r)} = 0$ ($\forall\, k \geqslant k_0$). 于是，按 Ch.-K. 方程，可得：

$$0 = p_{iu}^{(kd+r)} = \sum_{s \in K} p_{is}^{(kd+r-1)} \cdot p_{su}.$$

注意，若 $s \notin K_{r-1}$，则 $p_{is}^{(kd+r-1)} = 0$ ($\forall\, k \geqslant k_0$)(否则有 $s \in K_{r-1}$). 如果当 $s_0 \in K_{r-1}$ 时，$p_{s_0u} > 0$，则关于此 s_0 总可找到 $k \geqslant k_0$，使得 $p_{is_0}^{(kd+r-1)} > 0$. 故

$$0=\sum_{s\in K_{r-1}} p_{is}^{(kd+r-1)}\cdot p_{su}\geqslant p_{is_0}^{(kd+r-1)}\cdot p_{s_0u}>0.$$

这个矛盾表明：$p_{su}=0\ (\forall\ s\in K_{r-1},u\notin K_r)$.

$$\Rightarrow 1=\sum_{u\in K} p_{su}=\sum_{u\in K_r} p_{su}(\forall\ s\in K_{r-1}).$$

即 K_m 经一步即可达 K_{m+1}.

下面证明定理中的最后结论. 为此，先证

$$\sum_{u\in K_m} p_{uj}>0\ (\forall\ j\in K_{m+1}).$$

事实上，假如，$\exists\ u\in K_r$，其中 $r=r_0(\mathrm{mod}d)$，$m=m_0(\mathrm{mod}d)$，且 $r_0\neq m_0$，$\ni p_{uj}>0$. 则按 Ch.-K. 方程，有

$$p_{ij}^{(kd+r+1)}\geqslant p_{iu}^{(kd+r)}\cdot p_{uj}>0(\text{当 } k \text{ 足够大时}).$$

$$\Rightarrow j\in K_{r+1}.$$

但 $r_0+1\neq m_0+1$. 这就出现矛盾. 故

$$p_{uj}=0\ (\forall\ u\notin K_m,j\in K_{m+1}).$$

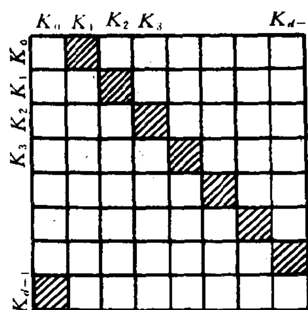
假如 $p_{uj}=0(\forall\ u\in K_m)$ 也成立，则相应于 K 的转移阵的第 j 列全为 0. 于是，j 为相对于 K 的非本质状态. 这与 K 的性质矛盾. 故对 $\forall\ j\in K_{m+1}$，$\exists\ u\in K_m$，$\ni p_{uj}>0$.

现在，利用这个结果来完成命题之证. 不妨设 $0\leqslant r_1<r_2\leqslant d-1$（$\because r_1=r_2$ 的情形结论显然真）. 按 Ch.-K. 方程，有

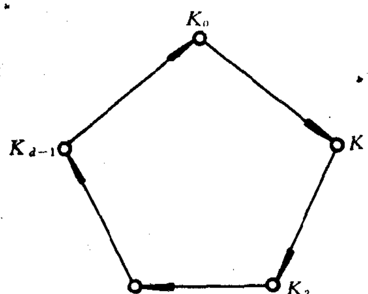
$$\begin{aligned}p_{j_1j_2}^{(r_2-r_1)}&=\sum_{i_l\in K_l,r_1<l<r_2} p_{j_1i_{r_1+1}}p_{i_{r_1+1}i_{r_1+2}}\cdots p_{i_{r_2-1}j_2}\\&=\sum_{i_l\in K_l,r_1<l<r_2} p_{j_1i_{r_1+1}}p_{i_{r_1+1}i_{r_1+2}}\cdots p_{i_{r_2-1}j_2}>0.\end{aligned}$$

$$\begin{aligned}p_{j_1j_2}^{(kd+r_2-r_1)}&=\sum_{s\in K} p_{j_1s}^{(kd)}p_{sj_2}^{(r_2-r_1)}\\&\geqslant p_{j_1j_1}^{(kd)}p_{j_1j_2}^{(r_2-r_1)}>0(\text{当 } k \text{ 足够大时}).\end{aligned}$$

注 4 假如 P_K 是相应于 K 的转移阵，$K=\bigcup_{r=0}^{d-1}K_r$，则 P_k 的结构如图 7.3-2 所示（分块形式）.



(a)



(b)

（空白处为 0）　　(a)块结构　　(b)运动图

图 7.3-2

定理 7.3.4　有限齐次 Markov 链的所有非常返状态组成的集合 K 不可能是闭集.

证　注意：集 K 是有限的. 按定理 7.2.1，系统从任意状态出发以概率 1 只能有限次返回 K 中的每个状态. 由 K 有限可得结论：系统从任意状态出发以概率 1 有限次返回集 K. 因此，系统最终必然离开 K，故 K 不能为闭集.

例 7.3.1　考察注 1 中的例，显然

$$p_{11}^{(n)} = \left(\frac{1}{2}\right)^n,\ p_{22}^{(n)} = \left(\frac{1}{2}\right)^n$$

$$\Rightarrow \sum_{n=1}^{\infty} p_{22}^{(n)} = 1,\ \sum_{n=1}^{\infty} p_{11}^{(n)} = 1.$$

由此可知，状态 1 和 2 都是非常返的. 从注 1 中也可看出：$\{1,2\}$ 不是 $E=\{0,1,2\}$ 的闭集.

推论 7.3.2　有限不可约 Markov 链的状态都是常返的.

证　按定理 7.3.4，此链的状态空间 E 不含非常返状态. 按引理 7.3.2，E 中状态全是常返的.

定理 7.3.5 有限齐次 Markov 链没有零常返状态.

证 设 E 为状态空间. 假定 $i\in E$ 为零常返状态. 则按定理 7.3.2,∃ 一个不可约常返闭集 C_l,∋ $i\in C_l$. 于是,$\sum_{j\in C_l}p_{ij}^{(n)}=1(\forall\, n\in\mathrm{N})$. C_l 为互通类,具有周期 d(含 $d=1$). 按状态判据(Ⅱ)(参看§7.2)和推论 7.2.1,有 $\lim\limits_{n\to\infty}p_{ij}^{(n)}=0$. 因 C_l 为有限集. 故 $\lim\limits_{n\to\infty}$ 与 $\sum\limits_{j\in C_l}$ 可以交换次序,这就得到:

$$1=\lim_{n\to\infty}\sum_{j\in C_l}p_{ij}^{(n)}=\sum_{j\in C_l}\lim_{n\to\infty}p_{ij}^{(n)}=0.$$

这个矛盾表明:E 不能含零常返状态.

定理 7.3.6 有限不可约 Markov 链的状态都是正常返的.

证 按推论 7.3.2,此链常返. 按定理 7.3.5,此链不含零常返状态,故结论真.

例 7.3.2 继续考察例 7.1.2(无限随机游动).

1. 所有状态都是互通的(由 $p_{ij}^{(n)}$ 的表示可见).

2. 所有状态都具有周期 2(由 $p_{ii}^{(2n+1)}=0, p_{ii}^{(2n)}>0$ 得出).

3. 当 $p\neq q$ 时,所有的状态都是非常返的.

事实上,$4p\cdot q<1$(当 $p\neq q$ 时),令 $t=4pq$,则按渐近公式,有

$$\sum_n p_{ii}^{(n)}=\sum_n p_{ii}^{(2n)}=\sum_n C_n\cdot\frac{t^n}{\sqrt{\pi}\,\sqrt{n}}\leqslant C\cdot\sum_n t^n<\infty.$$

按判据(Ⅲ),i 为非常返状态. $(\forall\, i\in E)$.

4. 当 $p=q$ 时,所有状态都是零常返的.

事实上,有

$$\overline{\lim_{n\to\infty}}p_{ii}^{(n)}=\lim_{n\to\infty}\frac{1}{\sqrt{\pi n}}=0.$$

$$\sum_n p_{ii}^{(n)}=\sum_n\frac{C_n}{\sqrt{\pi}}\cdot\frac{1}{\sqrt{n}}=+\infty.$$

按状态判据 2.(Ⅱ),i 为零常返状态$(\forall\, i\in E)$.

5. 无限随机游动是不可约 Markov 链(按定理 7.3.2).

例 7.3.3　离散分支过程.

1. 数学模型. 设 $\mathrm{N}=\{0,1,2,\cdots\}$, $E=\mathrm{N}$. ξ_n 表示群体第 n 代中个体的数目. 假定第 n 代中的第 k 个个体产生下一代的个体数为取值于 E 的 r.v. $Y_k^{(n)}$. 因此, 有 $\xi_{n+1}=\sum_{k=1}^{\xi_n} Y_k^{(n)} (\forall\, n\in\mathcal{N})$. 设 $Y=(Y_k, k\geqslant 1)$ 为独立同分布的 E-值 r.v. 列, 且 $P(Y_k=i)=p_i (\forall\, i\in E, k\geqslant 1)$. $(Y_k^{(n)}: k\geqslant 1, n\in\mathrm{N})$ 独立同分布, 且 $Y_k^{(n)}$ 与 Y_k 有相同的分布 $(\forall\, k\geqslant 1, n\in\mathrm{N})$. 按上述假设, 群体在时刻 n 之前的历史状况与之后的发展是独立的. 这表明: $\xi=(\xi_n, n\in\mathrm{N})$ 为一个 Markov 链. 现在给出此链的转移概率

$$P(\xi_{n+1}=j/\xi_n=i)=P\Big(\sum_{k=1}^{\xi_n} Y_k^{(n)}=j, \xi_n=i/\xi_n=i\Big)$$

$$=P\Big(\sum_{k=1}^{i} Y_k=j\Big).$$

令 $p_{ij}=P(\xi_{n+1}=j/\xi_n=i)$, 则 p_{ij} 与 $n\in\mathrm{N}$ 无关. 且

$$p_{ij}=\sum_{\substack{k_l\in E, 1\leqslant l\leqslant i\\ k_1+k_2+\cdots+k_i=j}} P(Y_l=k_l, 1\leqslant l\leqslant i)$$

$$=\sum_{\text{同前}} p_{k_1}p_{k_2}\cdots p_{k_i}. \tag{7.3-1}$$

综上分析 ξ 为齐次 Markov 链, 且转移阵由 (7.3-1) 式定义.

2. 母函数、定义函数

$$G(Z)=\sum_{k\in E} p_k Z^k,\quad F_n(Z)=\sum_{k\in E} p_{1k}^{(n)} Z^k (\forall\, n\in\mathrm{N}). \tag{7.3-2}$$

显然, G 和 F_n 在单位圆 $\{|Z|\leqslant 1\}$ 内解析, 且关于 Z 在原点处无穷次可微. 此外, 它们还具有下列性质:

G.1　$F_n(Z)=G^{(n)}(Z)$, 其中 $G^{(n)}(Z)=G(G^{(n-1)}(Z))=G\circ\cdots\circ G(Z)$.

事实上, 由 (7.3-1) 式可得

$$\sum_{k\in E} p_{ik}Z^k = \sum_{k\in E} Z^k \cdot \sum\nolimits^k p_{k_1}\cdots p_{k_i} = \Big(\sum_{k\in E} p_k Z^k\Big)^i = (G(Z))^i,$$

(7.3-3)

其中$\sum^k$对$k_1,\cdots,k_i$求和，而k_l满足约束：

$$k_l \in E,\ k_1 + \cdots + k_i = k.$$

$$\begin{aligned} F_{n+1}(Z) &= \sum_{k\in E} p_{1k}^{(n+1)} Z^k = \sum_{k\in E}\sum_{i\in E} p_{1i}^{(n)} p_{ik} Z^k \\ &= \sum_{i\in E} p_{1i}^{(n)} \sum_{k\in E} p_{ik} Z^k = \sum_{i\in E} p_{1i}^{(n)} (G(Z))^i \\ &= F_n(G(Z)) = F_n \circ G(Z) \\ &\Rightarrow F_{n+1}(Z) = G\circ\cdots\circ G(Z)\ (\text{函数}G\text{复合}n+1\text{次}), \end{aligned}$$

G. 2 $\quad F_{in}(Z) = (F_n(Z))^i = (G^{(n)}(Z))^i.$ (7.3-4)

其中$i\in E, F_{in}(Z) = \sum_{k\in E} p_{ik}^{(n)} Z^k$.

事实上，按 Ch.-K. 方程及(7.3-2)式，有

$$\begin{aligned} F_{in}(Z) &= \sum_{j\in E} p_{ij}^{(n-1)} (G(Z))^j = F_{in-1}(G(Z)) = \cdots = F_{i0}(G^{(n)}(Z)) \\ &= \sum_{j\in E} p_{ij}^{(0)} \cdot (G^{(n)}(Z))^j = (G^{(n)}(Z))^i. \end{aligned}$$

3. 状态分类. 设$p_0>0$. 显然，$p_{00}=1$，且

$$p_{i0} = P(Y_1+\cdots+Y_i=0) = \prod_{l=1}^{i} P(Y_l=0) = p_0^i(\forall\, i\in E).$$

因此，0 为吸收状态. 设$0\neq i\in E$，则

$$P(\sigma_i=\infty/\xi_0=i) \geqslant P(\xi_1=0/\xi_0=i) = p_0^i > 0.$$

因此，E中任意不为 0 的状态都是非常返的. 设$p_u>0(u\in E, 0\leqslant u\leqslant N)$，其中$N$为某正整数，则

(i)$i\longleftrightarrow j$，其中$i,j\in E$，且$0<i,j\leqslant N$；

(ii)当$0<i\leqslant N$时，i是非周期的.

事实上，由(7.3-1)式知：

$$p_{ij} = \sum\nolimits^j p_{k_1}p_{k_2}\cdots p_{k_i} \geqslant p_0^{i-1}p_j > 0.$$

这里求和符号$\sum^j$表示对$k_1,\cdots,k_i$求和，其中$k_i\in E$，且

$$k_1 + \cdots + k_i = j.$$

由此及对称性即可得(i). 为证(ii),考虑下列各式:

$$p_{1u} = p_u, p_{u1} = u p_0^{u-1} p_1 (\forall\, u \in E).$$

$$p_{11}^{(2)} = \sum_{u \in E} u p_0^{u-1} p_1 \cdot p_u \geqslant p_1^2 > 0.$$

因此,状态 1 非周期. 同理可得状态 2 非周期. 由此,即可归纳出(ii).

4. 状态空间 E 的分解

C.1　当 $P(Y_k=0)=p_0>0$ 时,$K=\{0\}$,$I=\{1,2,\cdots\}$,$E=K+I$. 其中 K 为常返闭集. I 为非常返状态集(不是闭集).

C.2　当 $p_0=1$ 时,$p_{i0}=1(\forall\, i\in E)$,

$$p_{ij} = 0\ (\forall\, j \in E, j \neq 0, i \in E).$$

这时,K 为闭集. I 为非本质状态集,不是闭集.

C.3　当 $p_0=0$ 时,$p_{0i}=p_{i0}=0(\forall\, i\in E, i\neq 0)$,$p_{00}=1$. 这时,$K$ 为闭集,I 为闭集.

5. 灭种界限

L.1　$p_{i0}^{(n)}=(p_{10}^{(n)})^i=\zeta_n^i$,其中 $i\in E$,$n\in\mathscr{N}$,$\zeta_n=p_{10}^{(n)}$.

事实上,按(7.3-4)式,有

$$p_{i0}^{(n)} = F_{in}(0) = (F_n(0))^i = (p_{10}^{(n)})^i.$$

L.2　$\zeta=\lim\limits_{n\to\infty}\zeta_n$ 存在,且满足方程:$\zeta=G(\zeta)$.

事实上,$G(z)$在 $z\in[0,1]$上是单调非降的,且

$$\zeta_{n+1} = p_{10}^{(n+1)} = \sum_{k\in E} p_{1k} p_{k0}^{(n)} = \sum_{k\in E} p_k \cdot \zeta_n^k = G(\zeta_n).$$

这里 $p_{1k}=p_k$. 于是,$\zeta_2=G(\zeta_1)\geqslant p_0=\zeta_1$. 假如$(\zeta_n, n\in \mathrm{N})$具有单调性,则极限 ζ 存在,且满足方程:$\zeta=G(\zeta)$.

考察差值:

$$\begin{aligned}\zeta_{n+1} - \zeta_n &= G(\zeta_n) - G(\zeta_{n-1}) = E(\zeta_n^Y - \zeta_{n-1}^Y) \\ &= E\Big[(\zeta_n - \zeta_{n-1}) \sum_{l=0}^{Y-1} \zeta_n^l \zeta_{n-1}^{Y-1-l} \cdot 1_{\{Y>0\}}\Big]\end{aligned}$$

$$= (\zeta_n - \zeta_{n-1})E\left(\sum_{l=0}^{Y-1}\zeta_n^l\zeta_{n-1}^{Y-1-l}1_{\{Y>0\}}\right).$$

由此递推关系及$\zeta_2 \geqslant \zeta_1$立即可得:$\zeta_{n+1} \geqslant \zeta_n$.

L.3 ζ为方程$z=G(z)$的最小正根.($\zeta=\lim\limits_{n\to\infty}\zeta_n$).

事实上,设u为$z=G(z)$的任一正根.则

$$\zeta_1 = p_0 = G(0) \leqslant G(u) = u;$$

$$\zeta_2 = G(\zeta_1) \leqslant G(u) = u;$$

$$\Rightarrow \zeta_n \leqslant G(u) = u. \Rightarrow \zeta \leqslant u.$$

L.4 设$1>p_0>0$,$m=EY(Y=Y_k)$,$G(z)=Ez^Y$,则"方程$z=G(z)$在$(0,1)$中有解"等价于"$m>1$".

事实上,令$f(z)=z-G(z)$,则$f(0)=-p_0<0$.

$$f'(z) = 1 - E(Yz^{Y-1}) > 0 \text{ 于 } z = 0 \text{ 附近.}$$

$$\Rightarrow f(z)\uparrow \text{ 于 } z = 0 \text{ 的附近.}$$

注意:$f(1)=1-G(1)=0$, $f'(1)=1-EY=1-m$.

若$m>1$,则$f(z)\downarrow 0$于$z=1$附近.故$z=G(z)$于$(0,1)$中有解.若$m<1$,则$f(z)$单调上升于$(0,1)$上(这是因为$f'(z)=1-EYz^{Y-1}\geqslant 1-m>0$于$(0,1)$上).这时,$z=G(z)$在$(0,1)$内无解.若$m=1$,则$f'(z)\geqslant 0$,且

$$f''(z) = -E(Y(Y-1)z^{Y-2}) < 0 \text{ 于}(0,1)\text{ 上.}$$

因此,$f(z)$严格单调上升到0.从而,$z=G(z)$于$(0,1)$内无解.

归纳上述分析,可得如下结论.

定理 7.3.7 假设种群的个体不产生下一代的概率$p_0\in(0,1)$.种群的个体产生下一代的平均值为m.则当$m\leqslant 1$时,该群体以概率1迟早消失(即灭种);当$m>1$时,该群体迟早消失的概率为ζ^i,无限发展的概率为$1-\zeta^i$,其中$i\in E$表示种群开始的成员数,ζ为方程$z=G(z)$在$(0,1)$内的最小正根.

证 当$m\leqslant 1$时,方程$z=G(z)$于$(0,1)$内无解.而ζ为满足此方程的最小正解.因此,$\zeta=1$,即

$$\lim_{n\to\infty} p_{i0}^{(n)} = \lim_{n\to\infty}\zeta_n^i = 1 \ (\forall\, i \in E).$$

这表明：随着时间的推移，群体自任意状态出发都终归消亡. 当 $m>1$ 时，有

$$\lim_{n\to\infty} p_{i0}^{(n)} = (\lim_{n\to\infty}\zeta_n)^i = \zeta^i (\forall\ i \in E).$$

此极限表明：群体开始时有成员 i 个，终归消亡的概率为 ζ^i. 注意，当 $m>1$ 时，方程 $Z=G(Z)$ 有解于 $(0;1)$ 中. 因此，$0<\zeta<1$. 由此可见，随着群体的初态增加，消亡的概率愈来愈小.

§7.4 $p_{ij}^{(n)}$的渐近性与平稳分布

前两节介绍状态分类和状态空间的分解. 并强调指出：不可约 Markov 链在 Markov 理论中是最基本的类型之一. 本节主要涉及这一类型，并讨论平稳 Markov 链的特性. 这里总假定 E 是不可约的. 按引理 7.3.2，E 要么全常返，要么全非常返，如何判断一个不可约 Markov 链是常返的问题与以下方程的性质有密切关系.

设 $\mathscr{P}$ 为转移阵. $x=(x_j, j\in E)^T$（"T"表示转置）. 则由 $\mathscr{P}$ 可以产生下列方程组：

$$x^T = x^T\mathscr{P} \text{ 或} \sum_{j\in E} p_{ji}x_j = x_i (\forall\ i \in E); \qquad (7.4\text{-}1)$$

$$x = \mathscr{P}x \text{ 或} \sum_{j\in E} p_{ij}x_j = x_i (\forall\ i \in E); \qquad (7.4\text{-}2)$$

$$x \geqslant \mathscr{P}x \text{ 或} \sum_{j\in E} p_{ij}x_j \leqslant x_i (\forall\ i \in E). \qquad (7.4\text{-}3)$$

定理7.4.1 不可约 Markov 链常返的充要条件是：不等式组 (7.4-3) 除 $x_i=C$（常数）外，没有任何非负的非平凡解.

证 必要性. 设链 ξ 不可约常返. 假如 (7.4-3) 有不全为 0 的非负解 x，如分量 $x_l>0$. 不可约性表明：$i\longleftrightarrow l$. 于是，$\exists\ n\geqslant 1$，$\ni$ $p_{il}^{(n)}>0$. 由方程 (7.4-3) 及 Ch.-K. 方程可推出：

$$x_i \geqslant \sum_{j\in E} p_{ij}^{(m)} x_j (\forall\ i \in E, m \in \mathscr{N})$$

$$\Rightarrow x_i \geqslant p_{il}^{(n)} x_l > 0\ (\forall\ i \in E).$$

即 x 的所有分量为正数. 现在,任取状态 $l_0\in E$. 令 $y_i=\frac{x_i}{x_{l_0}}$,则

$$y_i\geqslant\sum_{j\in E}p_{ij}y_j=p_{il_0}+\sum_{j\neq l_0}p_{ij}y_j.$$

叠代一次得

$$y_i\geqslant p_{il_0}+\sum_{j\neq l_0}p_{ij}p_{jl_0}+\sum_{j\neq l_0}\sum_{k\neq l_0}p_{ij}p_{jk}y_k. \tag{7.4-4}$$

按定理 7.2.3(i),有

$$p_{il_0}=f_{il_0}^{(1)},p_{il_0}^{(2)}=f_{il_0}^{(1)}p_{l_0l_0}+f_{il_0}^{(2)}$$

$$\Rightarrow f_{il_0}^{(2)}=p_{il_0}^{(2)}-f_{il_0}^{(1)}p_{l_0l_0}=\sum_{j\neq l_0}p_{ij}\cdot p_{jl_0}.$$

定义符号:${}_{l_0}p_{ik}^{(2)}=\sum\limits_{j\neq l_0}p_{ij}p_{jk}$,并代入(7.4-4)式,得

$$y_i\geqslant f_{il_0}^{(1)}+f_{il_0}^{(2)}+\sum_{k\neq l_0}{}_{l_0}p_{ik}^{(2)}y_k.$$

采用归纳法,可得

$$y_i\geqslant\sum_{n=1}^{N}f_{il_0}^{(1)}+\sum_{k\neq l_0}{}_{l_0}p_{ik}^{(N)}\cdot y_k\geqslant\sum_{n=1}^{N}f_{il_0}^{(n)}(\forall\ N\in\mathrm{N}). \tag{7.4-5}$$

按 ξ 的常返性及定理 7.2.2,有 $f_{il_0}=1$. 故在(7.4-5)式中让 $N\to\infty$,可得:$y_i\geqslant f_{il_0}=1$ $(\forall\ i\in E)$. 从而

$$x_i\geqslant x_{l_0}(\forall\ i\in E,l_0\in E).$$

这表明:$x_i\equiv C$(常数). 这就证实:非负非平凡解只能为常数.

充分性:采用反证法. 假定 ξ 有一个非常返状态 l,即 $f_{ll}=1$,则至少存在一个 $i_0\in E(i_0\neq l)$,$\ni f_{i_0l}<1$. 事实上,若不然,则 $f_{il}=1$ $(\forall\ i\in E,i\neq l)$,按引理 7.2.3(ii),有

$$f_{ll}^{(n)}=\sum_{j\neq l}p_{lj}f_{jl}^{(n-1)}$$

$$\Rightarrow f_{ll}=\sum_{n=1}^{\infty}f_{ll}^{(n)}=f_{ll}^{(1)}+\sum_{n=2}^{\infty}\sum_{j\neq l}p_{lj}f_{jl}^{(n-1)}$$

$$= p_{ll} + \sum_{j \neq l} p_{lj} f_{jl} = 1.$$

这与 $f_{ll} < 1$ 矛盾. 类似地作法可证

$$f_{il} = p_{il} + \sum_{k \neq l} p_{ik} f_{kl} (\forall\, i \in E). \qquad (7.4\text{-}6)$$

现在,令 $x_l = 1, x_i = f_{il} (i \neq l)$,则由(7.4-5)式可知:方程组(7.4-3)有一个非常数非负非平凡解.($\because x_l = 1, x_{i_0} < 1$). 这与条件不符,故 $f_{ll} = 1$. 从而,l 为常返状态,即 ξ 为常返链.

下面讨论方程组(7.4-1)的可解性. 假如求和符号可以交换. 将方程组(7.4-1)叠代 n 次,并采用 Ch.-K. 方程,立即可得

$$x^T = x^T \mathscr{P}^n (\forall\, n \in \mathrm{N}). \qquad (7.4\text{-}7)$$

令 $Q_n = \frac{1}{n} \mathscr{P}^n$. 假如 $\lim_{n \to \infty} Q_n$ 存在,且为 Q,则

$$x^T = x^T Q.$$

为保证此处理过程能通过,对 x 需加某些限制. 这就引出了如下结论.

定理 7.4.2 设 ξ 为任意常返不可约 Markov 链,则方程(7.4-1)最多只有一个绝对可和解 x,满足要求

$$\sum_{i \in E} x_i = 1.$$

若 ξ 正常返,则方程(7.4-1)满足上述要求的解具有如下形式:

$$x_i = v_i = \lim_{n \to \infty} \frac{1}{N} \sum_{n=1}^{N} p_{ji}^{(n)} = \frac{1}{\mu_i} (\forall\, i, j \in E).$$

若 ξ 零常返,则方程(7.4-1)的绝对可和解只能是平凡解.

证 假如方程(7.4-1)存在绝对可和解 x,则由于求和次序可以交换而有

$$x^T = x^T \mathscr{P}^n \text{ 或 } x^T = x^T Q_n,$$

其中 $\qquad Q_n = \frac{1}{n} \mathscr{P}^n = [q_{ij}^{(n)}]_{i,j \in E}, \; q_{ij}^{(n)} = \frac{1}{n} \sum_{k=1}^{n} p_{ij}^{(k)}.$

由 ξ 的不可约常返性知:ξ 具有周期 d(含 $d=1$),按引理 7.2.5 及

定理 7.2.5，有

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N}p_{ii}^{(n)}=\lim_{N\to\infty}\frac{1}{Nd}\sum_{n=0}^{N}p_{ii}^{(nd)}=\frac{1}{\mu_i}.$$

利用关系式：

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}f_{ji}^{(n)}\cdot n=0,$$

并仿引理 7.2.5 之证，便可得

$$\lim_{N\to\infty}q_{ji}^{(N)}=\lim_{N\to\infty}\sum_{k=1}^{N}\frac{1}{N}f_{ji}^{(k)}\sum_{n=0}^{N-k}p_{ii}^{(n)}=f_{ji}\cdot\frac{1}{\mu_i}.$$

由 ξ 的不可约常返性及定理 7.2.2 知：$f_{ji}=1$. 故

$$v_i=\lim_{N\to\infty}q_{ji}^{(N)}=\frac{1}{\mu_i}\ (\forall\ i,j\in E).$$

按假定：x 可和，即 $\sum\limits_{i\in E}|x_i|<\infty$. 因此，下面的求极限过程是合理的.

$$x^T=\lim_{N\to\infty}x^TQ_N=x^T\lim_{N\to\infty}Q_N=x^TQ,.$$

其中
$$Q=[q_{ij}]_{i,j\in E},\ q_{ji}=\frac{1}{\mu_i}\ (\forall\ i,j\in E)$$

$$\Rightarrow x_i=v_i\sum_{j\in E}x_j(\forall\ i\in E).$$

假如方程(7.4-1)有两个绝对可和解 $x,\widetilde{x}$ 都满足定理中的条件，则

$$\sum_{i\in E}(x_i-\widetilde{x}_i)=0.$$

这时，$x-\widetilde{x}$ 也是方程(7.4-1)的绝对可和解，且分量之和为 0. 按前面的结果，应有

$$x_i-\widetilde{x}_i=v_i\sum_{j\in E}(x_j-\widetilde{x}_j)=0\ (i\in E).$$

这表明：满足定理中要求的绝对可和解是唯一的.

现在，设 ξ 是不可约正常返链. 则有 $\frac{1}{\mu_i}>0\ (\forall\ i\in E)$. 若要求 x 满足：$\sum\limits_{i\in E}x_i=1$，则按前面的结果，应有 $x_i=v_i=\frac{1}{\mu_i}\ (\forall\ i\in E)$. 因

此，对 x 的要求转化为：$\sum_{i\in E} v_i=1$. 为证明此式，设 K 为 E 中的任意有限子集，则由 Ch.-K. 方程，有

$$q_{si}^{(N+1)}=\frac{1}{N+1}\sum_{n=1}^{N+1}p_{si}^{(n)}=\frac{1}{N+1}p_{si}+\frac{1}{N+1}\sum_{n=2}^{N+1}\sum_{j\in E}p_{sj}^{(n-1)}p_{ji}$$

$$\geqslant\frac{1}{N+1}p_{si}+\frac{N}{N+1}\sum_{j\in K}q_{sj}^{(N)}p_{ji}$$

$$\Rightarrow q_{si}\geqslant\sum_{j\in K}q_{sj}p_{ji}(\forall\ E\text{ 中的有限子集 }K)$$

$$\Rightarrow q_{si}\geqslant\sum_{j\in E}q_{sj}p_{ji}\text{，即 }v_i\geqslant\sum_{j\in E}v_jp_{ji}.$$

反复利用这种处理格式，便可得

$$v_i\geqslant\sum_{j\in E}v_jp_{ji}^{(n)}(\forall\ n\geqslant 1,i\in E).$$

现在，我们要证明此式的等号成立，假若不然，则

$$\exists\ n\geqslant 1\text{ 及 }l\in E,\ \ni v_l>\sum_{j\in E}v_jp_{jl}^{(n)}.$$

让 K 为 E 中的任意有限子集，则

$$\sum_{i\in K}v_i=\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}\sum_{i\in K}p_{ji}^{(n)}\leqslant 1\Rightarrow\sum_{i\in E}v_i\leqslant 1.$$

这表明：v 是非负可和的，于是，下面的求和号交换是合理的：

$$\sum_{i\in E}v_i>\sum_{i\in E}\sum_{j\in E}v_jp_{ji}^{(n)}=\sum_{j\in E}v_j.$$

这个矛盾表明：

$$v_i=\sum_{j\in E}v_jp_{ji}^{(n)}(\forall\ n\geqslant 1,i\in E)$$

$$\Rightarrow v_i=\sum_{j\in E}v_jq_{ji}^{(n)}$$

$$\Rightarrow v_i=\lim_{n\to\infty}\sum_{j\in E}v_j\cdot q_{ji}^{(n)}=\sum_{j\in E}v_j\cdot\lim_{n\to\infty}q_{ji}^{(n)}=v_i\sum_{j\in E}v_j.$$

（这里用到：v 是非负可和的性质）. 因正常返导致 $v_i>0(\forall\ i\in E)$. 故 $\sum_{j\in E}v_j=1$.

最后，设 ξ 是不可约零常返的. 则 $\mu_i=\infty(\forall\ i\in E)$. 从而 $v_i=$

0. 于是，$x^T = x^T Q = 0$. 由绝对可和解的唯一性，此解是唯一的平凡解.

注 1 零常返的不可约 Markov 链在绝对可和解范围内只有平凡解. 假如超出此范围，它是否存在非负非平凡解呢？为了回答这个问题，需要考虑 taboo 概率.

推论 7.4.1 不可约 Markov 链正常返的充要条件是：方程(7.4.1)存在唯一非负非平凡的绝对可和解 x，于 $\sum_{i\in E} x_i = 1$，且 $x_i = v_i = \frac{1}{\mu_i} (\forall\ i \in E)$.

推论 7.4.2 不可约 Markov 链是遍历的充要条件是：方程(7.4-1)存在唯一非负非平凡的绝对可和解 x，使得：$\sum_{i\in E} x_i = 1$，且

$$x_i = v_i = \lim_{n\to\infty} p_{ji}^{(n)} = \frac{1}{\mu_i} \ (\forall\ i, j \in E).$$

现在，开始介绍 taboo 概率. 记

$$ {}_l p_{ij}^{(n)} = \sum_{\substack{j_r \neq l,\\ 1\leqslant r\leqslant n-1}} p_{ij_1} p_{j_1 j_2} \cdots p_{j_{n-1} j}.$$

用以表示“系统从状态 i 出发在时刻 n 到达状态 j，但中途不访问状态 l”的概率(前面已见过).

定义 7.4.1 (taboo 概率)设 $H \subset E$. 令

$$\begin{aligned} {}_H p_{ij}^{(n)} &= P(\{\text{系统在时刻 } n \text{ 命中状态 } j\text{，但中途不进入 } H\} / \xi_0 = i) \\ &= P(\xi_r \notin H, 1 \leqslant r \leqslant n-1; \xi_n = j / \xi_0 = i). \end{aligned}$$

称 $\{{}_H p_{ij}^{(n)}: n \in \mathrm{N}, i、j \in E\}$ 为具有禁固集 H 的 taboo 概率. 若 $H = \{l\}$，则简记 ${}_H p_{ij}^{(n)} = {}_l p_{ij}^{(n)}$.

显然，${}_l p_{ij}^{(1)} = p_{ij}$，${}_j p_{ij}^{(n)} = f_{ij}^{(n)}$，${}_i p_{ij}^{(0)} = p_{ij}^{(0)}$.

下面介绍 taboo 概率的某些性质.

T.1 ${}_l p_{ij}^{(n)} = \sum_{k=1}^{n} {}_l f_{ij}^{(k)} \cdot {}_l p_{jj}^{(n-k)}$，其中 ${}_l f_{ij}^{(k)} = {}_{\{l,j\}} p_{ij}^{(k)} \ (k \geqslant 0)$；

$${}_l p_{ij}^{(n+1)} = \sum_{s\neq l} {}_l p_{is}^{(n)} p_{sj}.$$

事实上，第一式在形式上与引理 7.2.3(i)一致，其推导也类似. 第二式由定义而得.

T. 2
$$ {}_l p_{ij}^{(n)} = \sum_{k=0}^{n} {}_l p_{ii}^{(k)} \cdot {}_{\{i,l\}} p_{ij}^{(n-k)} (i \neq j). $$
$$ f_{ij}^{(n)} = \sum_{k=0}^{n} {}_j p_{ii}^{(k)} \cdot {}_i f_{ij}^{(n-k)} (i \neq j). $$

事实上，令 ${}_l A_j^{(k)} = \{\xi_r \neq l, 1 \leqslant r \leqslant k-1; \xi_k = j\}$，
$$ B(n,k) = \{\xi_k = i, \xi_r \neq i, k+1 \leqslant r \leqslant n\}. $$
则第一式等价于下式：
$$ \{\xi_0 = i\} \cap {}_l A_j^{(n)} = \bigcup_{k=0}^{n} \{\xi_0 = i\} \cap {}_l A_j^{(k)} \cap B(n,k) \quad (7.4\text{-}8) $$
定义 $\bar{\xi}_k = \xi_{n-k}, 0 \leqslant k \leqslant n$，
$$ \tau_i = \inf\{0 \leqslant k \leqslant n : \bar{\xi}_k = i\}, \ \inf\{\varnothing\} = n+1. $$
则 $P(\tau_i \leqslant n+1) = 1$，且 $(\tau_i = k) = B(n,k)$ $(0 \leqslant k \leqslant n)$.
$$ (\tau_i = n+1) = \{\bar{\xi}_k \neq i, 0 \leqslant k \leqslant n\} = \{\xi_k \neq i, 0 \leqslant k \leqslant n\}. $$
不难看出
$$ \{\xi_0 = i\} \cap {}_l A_j^{(k)} \cap (\tau_i = n+1) = \varnothing. $$
故(7.4-7)式成立.

在第一式中，取 $l = j$，即得第二式.

T. 3 母函数：${}_l P_{ij}(Z) = \sum_{n=0}^{\infty} {}_l p_{ij}^{(n)} Z^n$，
$$ {}_l F_{ij}(Z) = \sum_{n=0}^{\infty} {}_l f_{ij}^{(n)} Z^n. $$
具有下列关系：
$$ {}_l P_{ij}(Z) = {}_l F_{ij}(Z) \cdot {}_l P_{jj}(Z); \ F_{ij}(Z) = {}_j P_{ii}(Z) \cdot {}_i F_{jj}(Z). $$
事实上
$$ {}_l P_{ij}(Z) = \sum_{n=0}^{\infty} Z^n \cdot \sum_{k=0}^{n} {}_l f_{ij}^{(k)} \cdot {}_l p_{jj}^{(n-k)} $$
$$ = \sum_{k=0}^{\infty} {}_l f_{ij}^{(k)} \cdot \sum_{n=k}^{\infty} {}_l p_{jj}^{(n-k)} \cdot Z^n $$

$$= {}_lF_{ij}(Z)\cdot {}_lP_{jj}(Z).$$

$$F_{ij}(Z) = \sum_{n=0}^{\infty} f_{ij}^{(n)} Z^n = \sum_{n=0}^{\infty} {}_jp_{ii}^{(k)} \cdot {}_{\{i,j\}}p_{ij}^{(n-k)} Z^n$$

$$= \sum_{k=0}^{\infty} {}_jp_{ii}^{(k)} \sum_{n=k}^{\infty} {}_{\{i,j\}}p_{ij}^{(n-k)} Z^n$$

$$= {}_jP_{ii}(Z)\cdot\left(\sum_{n=0}^{\infty} {}_{\{i,j\}}p_{ij}^{(n)} Z^n\right)$$

$$= {}_jP_{ii}(Z)\cdot {}_iF_{ij}(Z).$$

T. 4　$i \longleftrightarrow j \Rightarrow {}_iF_{ij}(1) > 0,\ {}_jP_{ii}(1) = \dfrac{F_{ij}(1)}{{}_iF_{ij}(1)}.$

事实上,按引理 7.2.4,有 $f_{ij}>0$. 因此

$$\exists\, n \geqslant 1,\ \ni f_{ij}^{(n)} > 0.$$

从而,当 $i\neq j$ 时,按 T.2,有

$$\exists\, 1 \leqslant m \leqslant n,\ \ni {}_if_{ij}^{(m)} > 0.$$

当 $i=j$ 时,按 T.1,有

$$\exists\, 1 \leqslant m \leqslant n,\ \ni {}_if_{ii}^{(m)} > 0.$$

故${}_iF_{ij}(1)>0$. 利用 T.3 中的第二式,即得这里的结果.

T. 5　令${}_lG_{ij} = \sum_{n=0}^{\infty} {}_lp_{ij}^{(n)} = {}_lP_{ij}(1)$,则当 $i\to j$ 时,有

$$ {}_jG_{ii} = \frac{F_{ij}(1)}{{}_iF_{ij}(1)} < \infty;$$

$$ {}_iG_{ij} = {}_iF_{ij}(1)\cdot {}_iP_{jj}(1) = {}_iF_{ij}(1)\cdot {}_iG_{jj};$$

$$ {}_iG_{ij} \leqslant {}_iG_{jj} < \infty.$$

事实上,第一式即 T.4. 第二式为 T.3 中之第一式在 $Z=1$ 时的情形. 第三式由第二式推出.

定理 7.4.3　设 ξ 为常返不可约 Markov 链,则方程(7.4-1)允许有如下非负解 x:

$$x_l = 1,\ x_i = {}_iG_{li}\quad (i\neq l, i\in E),$$

其中 $l\in E$ 任意指定.

证 假定 $x_l=1, x_i={}_lG_{li}(i\neq l)$，则

$$\sum_{j\in E}x_jp_{ji}=p_{li}+{}_l\sum_{j\neq l}G_{lj}p_{ji}$$

$$=p_{li}+\sum_{j\neq l}{}_l\sum_{n=0}^{\infty}p_{lj}^{(n)}p_{ji}$$

$$=p_{li}+\sum_{n=1}^{\infty}{}_l\sum_{j\neq l}p_{lj}^{(n)}p_{ji}$$

$$=p_{li}+{}_l\sum_{n=1}^{\infty}p_{li}^{(n+1)}$$

$$={}_l\sum_{n=1}^{\infty}p_{li}^{(n)}={}_lG_{li}=x_i(i\neq l).$$

(这里用到 $\sum_{j\neq l}$ 与 $\sum_{n=0}^{\infty}$ 可以交换次序，易证这一点).

$$\sum_{j\in E}x_jp_{jl}=p_{ll}+{}_l\sum_{n=1}^{\infty}p_{ll}^{(n+1)}=\sum_{n=1}^{\infty}f_{ll}^{(n)}=f_{ll}=1.$$

综上分析，得 x 满足方程(7.4-1)之结论.

上述定理解决当指定 $x_l=1$ 时，方程(7.4-1)有解存在的问题. 现在考虑唯一性问题. 这里需要借助于逆 Markov 链的概念. 在介绍这一概念之前，先作一些解释.

引理 7.4.1 假定 ξ 是正常返不可约 Markov 链. 又设 $v=(v_i,i\in E)$ 为 ξ 的初始分布，这里的 v 按定理 7.4.2 中定义，即 $v_i=\frac{1}{\mu_i}>0(\forall\ i\in E)$，且 $\sum_{i\in E}v_i=1$. 这就是说，ξ 由 v 和 P 决定，考察 ξ 的一维分布：

$$P(\xi_n=j)=\sum_{i\in E}P(\xi_n=j/\xi_{n-1}=i)P(\xi_{n-1}=i)$$

$$=\sum_{i\in E}P(\xi_{n-1}=i)p_{ij};$$

$$P(\xi_1=j)=\sum_{i\in E}P(\xi_0=i)p_{ij}=\sum_{i\in E}v_ip_{ij}=v_j;$$

$$P(\xi_2=j)=\sum_{i\in E}P(\xi_1=i)p_{ij}=\sum_{i\in E}v_ip_{ij}=v_j;$$

$$\Rightarrow P(\xi_n = j) = v_j(\forall\ j \in E, n \in \mathrm{N}).$$

由此可见，ξ 的一维分布具有平稳性. 从而，ξ 为平稳不可约 Markov 链. 现在，定义 $q_{ij}^{(k)}=P(\xi_n=j/\xi_{n+k}=i)$，则

$$q_{ij}^{(k)} = P(\xi_{n+k} = i/\xi_n = j)\cdot\frac{P(\xi_n = j)}{P(\xi_{n+k} = i)} = p_{ji}^{(k)}\cdot\frac{v_j}{v_i}.$$

下面验证$(q_{ij}^{(k)}: k\in\mathscr{N}, i、j\in E)$满足转移概率的条件：

q. 1 $q_{ij}^{(0)}=\begin{cases}1, i=j,\\ 0, i\neq j.\end{cases}$

q. 2 $\sum\limits_{j\in E} q_{ij}^{(k)}=\Big(\sum\limits_{j\in E} v_j p_{ji}^{(k)}\Big)\dfrac{1}{v_i}=\dfrac{v_i}{v_i}=1.\ (\forall\ k\in\mathrm{N}).$

q. 3 Ch. -K. 方程对 q_{ij}成立，即

$$\begin{aligned}\sum_{j\in E} q_{ij}^{(k)} q_{js}^{(l)} &= \sum_{j\in E} p_{ji}^{(k)}\frac{v_j}{v_i}\cdot\frac{v_s}{v_j} p_{sj}^{(l)}\\ &= \frac{v_s}{v_i}\sum_{j\in E} p_{sj}^{(l)} p_{ji}^{(k)} = \frac{v_s}{v_i} p_{si}^{(k+l)} = q_{is}^{(k+l)}.\end{aligned}$$

这就证实 q 满足转移概率的要求.

令 $Q=[q_{ij}]_{i,j\in E}$，则 Q 为转移阵(齐次的!). 现在，假定 $\xi^{(v)}$是以 v 为初始分布，Q 为转移阵的齐次 Markov 链. 则 $\xi^{(v)}$的一维分布也是平稳的. 事实上

$$P(\xi_0^{(v)} = i) = v_i(\forall\ i \in E)\ (\text{初始分布}).$$

$$\begin{aligned}P(\xi_1^{(v)} = j) &= \sum_{i\in E} P(\xi_0^{(v)} = i) q_{ij}\\ &= \sum_{i\in E} v_i q_{ij} = \sum_{i\in E} p_{ji}\frac{v_j}{v_i} v_i = v_j\end{aligned}$$

$$\Rightarrow P(\xi_n^{(v)} = j) = v_j(\forall\ n \in \mathrm{N}, j \in E).$$

这表明：$\xi^{(v)}$的一维分布是平稳的. 这样就由 v 和 $\mathscr{P}$ 得到两个平稳练 $\xi,\xi^{(v)}$. 从 Q 的定义中不难看出，$\xi^{(v)}$是 ξ 反序的结果，即 $\xi^{(v)}=\{\cdots,\xi_2,\xi_1,\xi_0\}$. 这种链称为逆链，但下面的一种情形，却不具有这种平稳链.

引理 7.4.2 假定 ξ 是零常返不可约链. 按定理 7.4.3，方程

(7.4-1)存在一个正解 x，不可和. 但每个分量有限. 定义

$$q_{ij}^{(k)}=p_{ji}^{(k)}\frac{x_j}{x_i}\ (\forall\, k\in \mathrm{N}, i、j\in E).$$

类似于引理 7.4.1 中的证明，可知($q_{ij}^{(k)}:k\in\mathrm{N},i、j\in E$)为转移概率. 令 $Q=[q_{ij}]_{i,j\in E}$，则 Q 为转移阵，按定理 7.4.2，ξ 不存在平稳分布. 因此，不能像引理 7.4.1 中的情形构作两个平稳链. 但可根据 Q 及任意指定的初始分布构作一个链 ξ^*. 然而，ξ^* 不能通过 ξ 反序得到. 不过，这两种情形有一个共同点：两个转移阵的结构形式完全相同.

注 1　按定理 7.3.6，有限不可约 Markov 链是正常返的. 因此，引理 7.4.2 中的情形不会出现.

定义 7.4.2　若方程(7.4-1)有一个正解 x. $Q=[q_{ij}]_{i,j\in E}$，其中 $q_{ij}=p_{ji}\frac{x_j}{x_i}$. 则称具有转移阵 Q 的齐次 Markov 链 ξ^* 为原练 ξ 的逆链.

引理 7.4.1　若 Markov 链 ξ 不可约，且正常返(零常返)，则逆链 ξ^* 亦不可约，且正常返(零常返).

证　按 Q 的定义，有 $q_{ii}=p_{ii}(\forall\ i\in E)$. 故结论真.

定理 7.4.4　设状态空间 E 不可约且常返，则对$\forall$ 指定的 $l\in E$，方程(7.4-1)满足条件：$x_l=1$ 的有界非负解 x 是唯一的，且

$$x_i={}_lG_{li}=\lim_{N\to\infty}\Big(\sum_{n=1}^{N}p_{li}^{(n)}\Big)\cdot\Big(\sum_{n=1}^{N}p_{ll}^{(n)}\Big)^{-1}(i\neq l).\quad (7.4\text{-}9)$$

证　按推论 7.4.1，在正常返的情形下，所有正解之间仅相差一个常数因子. 因此，满足条件：$x_l=1$ 的正解是唯一的. 按定理 7.4.3，在零常返的情形下，方程(7.4-1)具有一个符合要求的有界正解 x：

$$x_l=1,\ x_i={}_lG_{li}\quad (i\neq l,\ i\in E).$$

现假定另有一个有界非负解 $\tilde{x}$ 满足要求 $x_l=1$. 则令 $x_i^*=x_i-\tilde{x}_i$ ($\forall\ i\in E$)，有

$$x_l^* = 0,\ x_i^* = \sum_{j\in E} x_j^* p_{ji} \quad (\forall\, i \in E, i \neq l).$$

注意到 x^* 有界，并利用 Ch. -K. 方程，即可得

$$x_i^* = \sum_{j\in E} x_j^* p_{ji}^{(n)} (\forall\, n \in \mathcal{N}, i \in E).$$

假若$\exists\ i_0 \in E, \ni x_{i_0}^* > 0$(如果$<0$,就将 x^* 全改符号). 设 K 为 E 中的有限子集. 则由链的零常返和不可约性知

$$\overline{\lim_{n\to\infty}} \sum_{j\in K} x_j^* p_{ji}^{(n)} \cdot \frac{1}{n} = \sum_{j\in K} x_j^* \overline{\lim_{n\to\infty}} \frac{1}{n} p_{ji}^{(n)} = 0 \ (\forall\, i \in E)$$

$$\Rightarrow x_{i_0}^* \leqslant \overline{\lim_{n\to\infty}} \lim_{K\uparrow E} \sum_{j\in K} x_j^* \cdot p_{ji_0}^{(n)} \cdot \frac{1}{n} = 0.$$

这就导致矛盾. 故 $x^* = 0$. 从而，唯一.

最后证明极限关系式(7.4-9)成立. 构作逆练 ξ^*，按引理 7.4.1，ξ^* 亦不可约，且零常返. 按定理 7.2.2，$\mathcal{F}_{ij} = 1 (\forall\, i, j \in E)$. 这里"$\mathcal{F}$"表示 ξ^* 的相应量. 注意

$$\sum_n q_{jj}^{(n)} = \infty.$$

及引理 7.2.3(i)和引理 7.2.5，有

$$\lim_{N\to\infty} \Big(\sum_{n=1}^N q_{ij}^{(n)}\Big)\Big(\sum_{n=1}^N q_{jj}^{(n)}\Big)^{-1} = \lim_{N\to\infty}\Big(\sum_{k=0}^N q_{jj}^{(k)} \sum_{n=0}^{N-k} \mathcal{F}_{ij}^{(n)}\Big)\Big(\sum_{n=1}^N q_{jj}^{(n)}\Big)^{-1}$$

$$= \lim_{N\to\infty} \sum_{n=0}^N \mathcal{F}_{ij}^{(n)} = \mathcal{F}_{ij} = 1.$$

注意：$q_{ij}^{(n)} = p_{ji}^{(n)} \dfrac{x_j}{x_i}$. 故由上式可得

$$\lim_{N\to\infty} \Big(\sum_{n=1}^N p_{ij}^{(n)}\Big)\Big(\sum_{n=1}^N p_{jj}^{(n)}\Big)^{-1} = \lim_{N\to\infty}\Big(\frac{x_i}{x_j}\sum_{n=1}^N q_{ij}^{(n)}\Big)\Big(\sum_{n=1}^N q_{jj}^{(n)}\Big)^{-1}$$

$$= \frac{x_i}{x_j}\ (\forall\, i, j \in E).$$

由此即得(7.4-9)式成立.

§7.5 遍历性定理

按定义 7.2.5，非周期正常返齐次 Markov 链简称为遍历链.假如此链不可约，则按 §4 中的讨论，它具有一个平稳分布是正的且唯一.由 v 为初始分布和 $\mathscr{P}$ 为转移阵可决定一个平稳 Markov 链.现在讨论这种平稳链的遍历性.

定义 7.5.1 设 $\xi=(\xi(t),t\in T)$ 为定义在 $(\Omega,\mathscr{F},P)$ 上取值于可测空间 $(\mathscr{X},\mathscr{B})$ 的随机过程.若对 $\forall\ n\geqslant 1,(t_1,\cdots,t_n)\subset T$ 及使 $t+t_k\in T$ 的 t，有

$$P(\xi(t+t_k)\in B_k,1\leqslant k\leqslant n)=P(\xi(t_k)\in B_k,1\leqslant k\leqslant n)$$
$$(\forall\ B_k\in\mathscr{B},1\leqslant k\leqslant n),$$

则称 ξ 为(严)平稳过程.

注 1 如果 ξ 只要求对 $n=1,2$ 具有定义 7.5.1 中的性质，则一般称 ξ 为宽平稳过程，以下所涉及的平稳过程均按定义 7.5.1 理解.

现在，设 $T=\{\cdots,-2,-1,0,1,2,\cdots\}$,

$$\mathscr{X}^T=\{u=(u_n,n\in T):u_n\in\mathscr{X},n\in T\}.$$

并让 ξ 为具有参数集 T 的 $\mathscr{X}$-值随机序列，P_ξ 表示由 ξ 诱导的 $(\mathscr{X}^T,\mathscr{B}^T)$ 上的概率测度.

定义 7.5.2 设 θ 是 $(\mathscr{X}^T,\mathscr{B}^T)\to(\mathscr{X}^T,\mathscr{B}^T)$ 的可测变换.若 θ 由如下关系定义：

$$u'=\theta u\ (\forall\ u\in\mathscr{X}^T),\ u'_n=u_{n+1}(\forall\ u\in T),$$

则称 θ 为平移算子.$\bar{\theta}$ 表示 θ 的逆，即

$$\bar{\theta}u_{n+1}=u_n(\forall\ u\in\mathscr{X}^T,n\in T).$$

引理 7.5.1 $\mathscr{X}$-值 r.v. 列 ξ 平稳的充要条件是：对 $\mathscr{B}^T$ 中的任意柱集 C，有

$$P_\xi(C)=P_\xi(\theta(C)),\text{其中 }\theta(C)=\{u':u'=\theta u,u\in C\}.$$

证 设 $\theta^n=\theta\circ\theta\circ\cdots\circ\theta$($\theta$ 的 n 重复合映射),

$$C = \{u \in \mathscr{X}^T : u_{t_k} \in B_k, t_k \in T, 1 \leqslant k \leqslant n\}.$$

则当 $m \geqslant 1$ 且满足 $m + t_k \in T$ 时，有

$$\theta^m(C) = \{u \in \mathscr{X}^T : u_{t_k+m} \in B_k, t_k \in T, 1 \leqslant k \leqslant n\}.$$

显然，有

$$P_\xi(C) = P_\xi(\theta^m(C)) \Longleftrightarrow P_\xi(C) = P_\xi(\theta(C)).$$

定义 7.5.3 设 $(U, \mathscr{U}, \mu)$ 为一测度空间. S 为 $(U, \mathscr{U})$ 到自身的可测变换，若对 $\forall\ A \in \mathscr{U}$，有

$$\mu(S^{-1}(A)) = \mu(A),$$

其中 $S^{-1}(A) = \{u : Su = u' \in A\}$，则称 S 为保测变换.

若 $SS^{-1} = S^{-1}S = I$，则说 S 可逆，且其逆为 S^{-1}.

引理 7.5.2 设 $(U, \mathscr{U}, \mu)$ 为一个 σ-有限测度空间. S 为此空间上的保测变换. $f = (f(u) : u \in U)$ 为 $\mathscr{U}$-可测的非负函数，则

$$\int_{S^{-1}(D)} f(Su)\mu(\mathrm{d}u) = \int_D f(u)\mu(\mathrm{d}u) \quad (\forall\ D \in \mathscr{U}).$$

证 先考虑 $f(u) = 1_A(u)$，其中 $A \in \mathscr{U}$. 则

$$\int_D f(u)\mu(\mathrm{d}u) = \mu(AD),$$

$$\begin{aligned} f(Su) &= 1_A(Su) = 1_{S^{-1}(A)}(u) \\ &\Rightarrow \int_{S^{-1}(D)} f(Su)\mu(\mathrm{d}u) = \mu(S^{-1}(A) \cap S^{-1}(D)) \\ &= \mu(S^{-1}(A \cap D)) = \mu(A \cap D) \\ &= \int_D f(u)\mu(\mathrm{d}u). \end{aligned}$$

由此立得，引理对简单函数真.

至于一般的非负可测函数 f，可作出一个非负简单函数列 $(f_n, n \geqslant 1)$，使得 $0 \leqslant f_n(u) \uparrow f(u)$. 利用单调收敛定理，即可得引理真.

定义 7.5.4 设 $(a_k, 1 \leqslant k \leqslant n)$ 为实数列. p 为正整数. $S(k, k+i) = \sum_{l=k}^{k+i} a_l$. 若 $(S(k, k+i), 0 \leqslant i \leqslant p-1)$ 中至少有一个非负，则称 a_k

具有 p-标志.

引理 7.5.3 在实数列 $a=(a_k,k\geqslant 1)$ 中,具有 p-标志的元素之和非负.

证 设 $k_1=\inf\{k\geqslant 1: a_k$ 具有 p-标志$\}$,则按定义 7.5.4,$\exists\ 0\leqslant r\leqslant p-1$, 于 $S(k_1,k_1+r)\geqslant 0$. 设

$$r_1=\inf\{0\leqslant \nu\leqslant p-1: S(k_1,k_1+\nu)\geqslant 0\},$$

则当 $r_1\neq 0$ 时,有 $S(k_1,k_1+h)<0(0\leqslant h<r_1)$. 这样的数对 (k_1,r_1) 具有下列特性:

M.1 $S(k_1,k_1+r_1)\geqslant 0$,

M.2 $a_{k_1},a_{k_1+1},\cdots,a_{k_1+r_1}$ 都具有 p-标志.

事实上,若 $r_1=0$,则按 k_1 之定义,a_{k_1} 具有 p-标志. 若 $r_1\neq 0$,则 $a_{k_1}<0$. 按 M.1,有

$$a_{k_1+1}+a_{k_1+2}+\cdots+a_{k_1+1+r_1-1}\geqslant -a_{k_1}>0.$$

故 a_{k_1+1} 具有 p-标志. 依此类推,即得 M.2.

现在,以序列 $(a_{k_1+r_1+k},k\geqslant 1)$ 取代序列 a,则可得数对 (k_2,r_2),具有性质 M.1 和 M.2. 继续这种作法,即可得数对列 $\{(k_i,r_i),i\geqslant 1\}$. 令

$$a^{(i)}=\{a_{k_i},\cdots,a_{k_i+r_i}\},$$

则 $S(k_i,k_i+r_i)\geqslant 0$; $a\setminus\bigcup_{i=1}^{\infty}a^{(i)}$ 中不含具有 p-标志的元素,令 $S=\sum_{i=1}^{\infty}S(k_i,k_i+r_i)\geqslant 0$,则 S 为 a 中所有具有 p-标志的元素之和.

引理 7.5.4 设 $(U,\mathscr{U},\mu)$ 为 σ-有限测度空间. f 为 U 上的 μ-可积实函数. S 为 $(U,\mathscr{U})$ 上的关于 μ 的保测变换. $E_0=\bigcup_{n=1}^{\infty}\left\{u:\sum_{k=1}^{n}f(S^{k-1}u)\geqslant 0\right\}$,则

$$\int_{E_0}f(u)\mu(\mathrm{d}u)\geqslant 0.$$

证 令 $f_m(u)=f(u)\cdot 1_{A_m}(u)$, $A_m=\{u:|f(u)|\leqslant m\}$.

$$E_0^m=\bigcup_{n=1}^{\infty}\left\{u:\sum_{k=1}^{n}f_m(S^{k-1}u)\geqslant 0\right\}.$$

则 $E_0^m\uparrow E_0$，$1_{E_0^m}(u)\cdot f_m(u)\xrightarrow[m\to\infty]{}1_{E_0}(u)f(u)$. 这些关系可由 f 是 μ-可积之假设条件推出. 按控制收敛定理，有

$$\int_{E_0}f(u)\mu(\mathrm{d}u)=\lim_{m\to\infty}\int_{E_0^m}f_m(u)\mu(\mathrm{d}u). \qquad (7.5\text{-}1)$$

因此，若能证明

$$\int_{E_0^m}f_m(u)\mu(\mathrm{d}u)\geqslant 0, \qquad (7.5\text{-}2)$$

则(7.5-1)式表明：引理成立.

不失一般性，设 f 有界，μ 有限. 这时(7.5-2)式中让 m 足够大，就变成(7.5-1)式，令

$$D_k=\{u:f(S^k u)\text{ 具有 }p\text{- 标志}\},$$

则($D_k, k\geqslant 0$)具有下列性质：

D. 1 $D_0=\left\{u:\sup_{\nu\leqslant p}\left(\sum_{s=0}^{\nu-1}f(S^s u)\right)\geqslant 0\right\}$；

D. 2 $D_n=S^{-n}D_0$.

事实上，按定义 7.5.4，有

$$u\in D_n\Rightarrow S^n u\in D_0\Rightarrow D_n\subset S^{-n}(D_0).$$

反之，让 $u_0\in D_0$，令 $u=S^{-n}u_0$，则

$$0\leqslant\sup_{\nu\leqslant p}\left(\sum_{k=0}^{\nu-1}f(S^k u_0)\right)=\sup_{\nu\leqslant p}\left(\sum_{k=0}^{\nu-1}f(S^{n+k}u)\right)$$
$$\Rightarrow u\in D_n\Rightarrow S^{-n}(D_0)\subset D_n.$$

联合起来，即得 D. 2.

D. 3 $\int_{D_n}f(S^n u)\mu(\mathrm{d}u)=\int_{D_0}f(u)\mu(\mathrm{d}u)$ (按引理 7.5.2)；

D. 4 $\int_{D_0}f(u)\mu(\mathrm{d}u)\geqslant 0$.

事实上，对$\forall$ $n\geqslant 1$，按 D. 2，有

$$\int_{D_0}f(u)\mu(\mathrm{d}u)=\sum_{k=0}^{n}\frac{1}{n+1}\int_{D_k}f(S^k u)\mu(\mathrm{d}u)$$

$$= \int_U \frac{1}{n+1} \sum_{k=0}^{n} f(S^k u) \cdot 1_{D_k}(u) \mu(\mathrm{d}u).$$

令 $g_n(u) = \frac{1}{n+1} \sum_{k=0}^{n} f(S^k u) \cdot 1_{D_k}(u)$，则 $(g_n, n \geqslant 1)$ 关于 μ 是一致可积的. 从而，可选一子列 $(g_{n_i}, i \geqslant 1)$ $(\mathrm{mod}\,\mu)$ 收敛于某个有界函数 $g(u)$. 按控制收敛定理，有

$$\int_{D_0} f(u) \mu(\mathrm{d}u) = \int_U g(u) \mu(\mathrm{d}u).$$

现在证明 $g(u) \geqslant 0 (\mathrm{mod}\,\mu)$. 为此，设

$$S_p(u) = \sum_{k=0}^{\infty} f(S^k u) 1_{D_k}(u),$$

即是 $\{f(u), f(Su), f(S^2 u), \cdots\}$ 的具有 p-标志的元素之和. 按引理7.5.3，有 $S_p(u) \geqslant 0$. 若 $S_p(u) < \infty$，则显然有 $g(u) = 0$. 若 $S_p(u) = \infty$，且 $\mu(u: g(u) < -\varepsilon) > 0$ $(\varepsilon > 0)$.

则 $$\lim_{n \to \infty} \mu\left(u: \sum_{k=0}^{n} f(S^k u) 1_{D_k}(u) < -\varepsilon(n+1)\right) > 0.$$

然而，另有

$$\lim_{n \to \infty} \mu\left(u: \sum_{k=0}^{n} f(S^k u) 1_{D_k}(u) < -\varepsilon(n+1)\right)$$
$$= \mu(u: S_p(u) = -\infty) = 0.$$

这就出现矛盾. 故对 $\forall\ \varepsilon > 0$，有

$$\mu(u: g(u) < -\varepsilon) = 0 \Rightarrow \mu(u: g(u) < 0) = 0.$$

记 $D_0 = D_0^p$，则 $D_0^p \uparrow E_0$ 当 $p \uparrow \infty$ 时，于是，按控制收敛定理，有

$$\int_{E_0} f(u) \mu(\mathrm{d}u) = \lim_{p \to \infty} \int_{D_0^p} f(u) \mu(\mathrm{d}u) \geqslant 0.$$

定理 7.5.1 （极大遍历性定理） 设 $(U, \mathscr{U}, \mu)$ 为 σ-有限测度空间. f 为 U 上的 μ-可积函数. S 为 $(U, \mathscr{U})$ 上的保测变换. 令

$$E_\lambda = \bigcup_{n=1}^{\infty} \left\{u: \frac{1}{n} \sum_{k=1}^{n} f(S^{k-1} u) \geqslant \lambda\right\}$$

$$= \left\{u: \sup_{n \geqslant 1} \left(\frac{1}{n} \sum_{k=1}^{n} f(S^{k-1} u)\right) \geqslant \lambda\right\} \quad (\forall\ \lambda \in \mathrm{R}),$$

则 $$\int_{E_\lambda} f(u)\mu(\mathrm{d}u) \geqslant \lambda \cdot \mu(E_\lambda) \ (\forall\ \lambda \in \mathrm{R}).$$

证 事实上，令 $f^*(u)=f(u)-\lambda$，并以 f^* 代替引理 7.5.4 中的 f，则当 $\mu(E_\lambda)<\infty$时，按引理 7.5.4，立即得到所要的结论. 假如 $\mu(E_\lambda)=\infty$，如果 $\lambda<0$，则结论自然成立. 如果 $\lambda\geqslant 0$，则 μ 为 σ-有限可知

$$\exists\ E_n \uparrow U,\ \ni\ \mu(E_n) < +\infty\ (\forall\ n \geqslant 1).$$

按控制收敛定理及下式：

$$\int_{E_\lambda} f(u) 1_{E_n}(u)\mu(\mathrm{d}u) \geqslant \lambda\mu(E_\lambda \cap E_n)$$

便可推出所要的结论.

定理 7.5.2(Birkhoff-Khinchin 定理) 设$(U,\mathscr{U},\mu)$为σ-有限测度空间，S 为$(U,\mathscr{U})$上关于 μ 的保测变换. f 为 U 上的 μ-可积函数. 则

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} f(S^k u) = f^*(u)\ (\mathrm{mod}\mu)\ (u \in U)$$

且极限 f^* 是 S-不变的，即

$$f^*(Su) = f^*(u)\ (\mathrm{mod}\mu)\ (u \in U),$$

且 f^* 是 μ-可积的. 此外，当 $\mu(U)<\infty$时，有

$$\int_U f^*(u)\mu(\mathrm{d}u) = \int_U f(u)\mu(\mathrm{d}u).$$

证 不失一般性，设 f 是非负的. 令

$$g^*(u) = \overline{\lim_{n\to\infty}} \frac{1}{n}\sum_{k=0}^{n-1} f(S^k u),\ g_*(u) = \underline{\lim_{n\to\infty}} \frac{1}{n}\sum_{k=0}^{n-1} f(S^k u).$$

则命题中的极限存在等价于“$g^*=g_*\ (\mathrm{mod}\mu)$. ”等价于“$\mu(K_{\alpha\beta})=0\ (\forall\ \alpha、\beta\in Q=\{\mathrm{R}$ 中有理数全体$\})$，其中 $K_{\alpha\beta}=\{u: g^*(u)>\beta>\alpha>g_*(u)\}$”.

注意：

$$g^*(Su) = \overline{\lim_{n\to\infty}} \frac{1}{n}\sum_{k=0}^{n-1} f(S^{k+1} u)$$

$$= \overline{\lim_{n\to\infty}}\left(\frac{n+1}{n}\cdot\frac{1}{n+1}\sum_{k=0}^{n}f(S^k u) - \frac{1}{n}f(u)\right)$$

$$= g^*(u)\ (\mathrm{mod}\mu).$$

类似地也有 $g_*(Su)=g_*(u)\ (\mathrm{mod}\mu)$. 因此, g^* 和 g_* 都是 S-不变的. 从而, $K_{\alpha\beta}$ 为 S-不变集, 即

$$S^{-1}(K_{\alpha\beta}) = K_{\alpha\beta}(\mathrm{mod}\mu).$$

令 $K_\beta=\{u: g^*(u)>\beta\}$, 有

$$E_\beta = \left\{u: \sup_n\left(\frac{1}{n}\sum_{k=0}^{n-1}f(S^k u)\right)\geqslant\beta\right\},$$

则 $K_{\alpha\beta}\subset K_\beta\subset E_\beta$. 按定理 7.5.1, 有

$$\int_{K_{\alpha\beta}}f(u)\mu(\mathrm{d}u)\geqslant\beta\cdot\mu(K_{\alpha\beta}).$$

以 $-f$ 代替 f, 再利用此定理便可得

$$\int_{K_{\alpha\beta}}f(u)\mu(\mathrm{d}u)\leqslant\alpha\cdot\mu(K_{\alpha\beta}).$$

结合起来即得 $\mu(K_{\alpha\beta})=0\ (\forall\ \alpha,\beta\in Q)$

$$\Rightarrow g^*(u) = g_*(u)\ (\mathrm{mod}\mu).$$

令 $f^*(u)=g^*(u)$, 则 f^* 是 S-不变的. 为证 f^* 是 μ-可积的. 令 $A_{kn}=\left\{u: \frac{k}{2^n}\leqslant f^*(u)<\frac{k+1}{2^n}\right\}$, 则

$$U = \bigcup_{k=-\infty}^{\infty}A_{kn}(\forall\ n\geqslant 0)\ (\mathrm{mod}\mu);$$

$$S^{-1}A_{kn} = A_{kn}\subset E_{\frac{k}{2^n}}.$$

按定理 7.5.1, 有

$$\int_{A_{kn}}f(u)\mu(\mathrm{d}u)\geqslant\frac{k}{2^n}\mu(A_{kn}),$$

$$\int_{A_{kn}}f(u)\mu(\mathrm{d}u)\leqslant\frac{k+1}{2^n}\mu(A_{kn})$$

$$\Rightarrow\left|\int_{A_{kn}}(f(u)-f^*(u))\mu(\mathrm{d}u)\right|\leqslant 2^{-n}\mu(A_{kn}).$$

μ 的 σ-有限性表明：$\exists\ B_m$，$\ni U=\bigcup_{m=1}^{\infty}B_m$，且

$$\mu(B_m)<\infty\ (\forall\ m\geqslant 1).$$

利用上面的估计，得

$$\int_{B_m}f(u)\mu(\mathrm{d}u)=\int_{B_m}f^*(u)\mu(\mathrm{d}u)\ (\forall\ m\geqslant 1).$$

让 $m\to\infty$，即知 f^* 是 μ-可积的.

如果 $\mu(U)<\infty$，则取 $B_m=U\ (\forall\ m\geqslant 1)$，立即可得出定理中的最后等式.

注 1 引理 7.5.4 和定理 7.5.1 的结论可以加强，即对 A_λ 中的任意可测子集 A，均有

$$\int_A f(u)\mu(\mathrm{d}u)\geqslant 0\quad(\text{引理 7.5.4});$$

$$\int_A f(u)\mu(\mathrm{d}u)\geqslant\lambda\mu(A)\quad(\text{定理 7.5.1}).$$

这只要审查一下它们的证明即可看出.

上述各结论对可测集的 S-不变性很敏感. 下面给出某些有关的结果.

定义 7.5.4 设 $(U,\mathscr{U},\mu)$ 为 σ-有限测度空间. 若 $A\in\mathscr{U}$，且 $\mu(A\Delta S^{-1}(A))=0$（对称差的测度）. 则说 A 关于 μ 是 S-不变的. 记

$$\mathscr{G}_S=\{A\in\mathscr{U}:A\text{ 关于 }\mu\text{ 是 }S\text{-不变的}\},$$

则 $\mathscr{G}_S$ 为 $\mathscr{U}$ 中的 S-不变子 σ-代数.

推论 7.5.1 若在定理 5.2 中加设 $\mu(U)=1$. 则

$$f^*(u)=E(f/\mathscr{G}_S)\ (\mathrm{mod}\mu),$$

其中 $E(\cdot/\cdot)$ 为关于 μ 的条件期望算子.

证 按定理 7.5.2，有

$$f^*(Su)=f^*(u)\ (\mathrm{mod}\mu).$$

这里 f^* 按定理 7.5.2 构作. 因此，f^* 是 $\mathscr{G}_S$-可测的. 设 g 为 U 上的任意有界 $\mathscr{G}_S$-可测的函数. 显然

$$(g(\cdot)\cdot f(\cdot))^*(u)=\lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1}g(S^ku)\cdot f(S^ku)$$
$$=g(u)f^*(u)\ (\mathrm{mod}\mu).$$

这里用到 g 的 $\mathscr{G}_S$-可测性，即 $g(Su)=g(u)(\mathrm{mod}\mu)$. 按定理 7.5.2，得

$$E(g\cdot(f^*-f))=E((g(\cdot)\cdot f(\cdot))^*-g\cdot f)=0.$$

按 g 的任意性，即知：定义 3.2.1 的条件成立. 故

$$f^*=E(f/\mathscr{G}_S)\ (\mathrm{mod}\mu).$$

推论 7.5.2 若在定理 7.5.2 中加设 $\mu(U)<\infty$ 和 $E_\mu|f|^r<\infty(r\geqslant1)$，其中 E_μ 表示按测度 μ 求积分. 则

$$\lim_{n\to\infty}E_\mu\left|\frac{1}{n}\sum_{k=0}^{n-1}f(S^ku)-f^*(u)\right|^r=0.$$

证 令 $g=f-f^*$，则考虑到 f^* 的 S-不变性，有

$$\frac{1}{n}\sum_{k=0}^{n-1}f(S^ku)-f^*(u)=\frac{1}{n}\sum_{k=0}^{n-1}g(S^ku).$$

将 μ 概率化，不妨设 $\mu(U)=1$. 则按推论 7.5.1，得 $\|f^*\|_r\leqslant\|f\|_r$. 因此，$\|g\|_r<\infty$. 按 Minkowski 不等式及引理 7.5.2，有

$$\left\|\frac{1}{n}\sum_{k=0}^{n-1}g(S^ku)\right\|_r\leqslant\|g\|_r$$

$$\Rightarrow\frac{1}{n}\sum_{k=0}^{n-1}g(S^ku)\xrightarrow[n\to\infty]{}0\ (\mathrm{mod}\mu)\ (\text{按定理 7.5.2}).$$

采用 Lebesque 控制收敛定理，可得

$$\lim_{n\to\infty}\left\|\frac{1}{n}\sum_{k=0}^{n-1}g(S^ku)\right\|_r^r=\left\|\lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1}g(S^ku)\right\|_r^r=0.$$

推论 7.5.3 设 $\xi=(\xi_n,n\in T)$ 为定义在 $(\Omega,\mathscr{F},P)$ 上取值于 $(\mathscr{X},\mathscr{B})$ 的平稳序列. P_ξ 是由 ξ 诱导的 $(\mathscr{X}^T,\mathscr{B}^T)$ 上的概率测度. f 是 $(\mathscr{X}^m,\mathscr{B}^m)$ 上的实值可测函数，且 $E|f(\xi_0,\cdots,\xi_{m-1})|<\infty$. θ 为 $(\mathscr{X}^T,\mathscr{B}^T)$ 上的平移算子. $\mathscr{G}_\theta$ 是 $\mathscr{B}^T$ 中关于 P_ξ 的 θ-不变的子 σ-代数，即

$$P_\xi(B\Delta\theta^{-1}B)=0\ (\forall\ B\in\mathscr{G}_\theta),$$

则

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1}f(\xi_k,\cdots,\xi_{k+m-1})=E(f(\xi_0,\cdots,\xi_{m-1})/\widetilde{\mathscr{G}}_\theta)\ (\mathrm{mod}P),$$

其中 $\widetilde{\mathscr{G}}_\theta=\xi^{-1}(\widetilde{\mathscr{G}}_\theta)$ 为 $\mathscr{F}$ 中的子 σ-代数.

证 令 $\gamma(u)=(u_0,\cdots,u_{m-1}),u\in\mathscr{X}^T$,则

$$\gamma(\theta^{-k}u)=(u_k,\cdots,u_{k+m-1}).$$

令
$$\widetilde{f}(u)=f\circ\gamma(u)=f(u_0,\cdots,u_{m-1}),$$

则
$$\widetilde{f}(\theta^{-k}u)=f\circ\gamma(\theta^{-k}u)=f(u_k,\cdots,u_{k+m-1}).$$

由 ξ 的平稳性知:θ^{-1} 是 $(\mathscr{X}^T,\mathscr{G}_\theta,P_\xi)$ 上的保测变换. 按定理 7.5.2,有

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1}\widetilde{f}(\theta^{-k}u)=E(\widetilde{f}/\mathscr{G}_\theta)(\mathrm{mod}P_\xi).$$

将上式中的 u 改成 ξ,并注意 $E(\widetilde{f}/\mathscr{G}_\theta)(u)|_{u=\xi}$ 是 $\mathscr{G}_\theta$-可测的,则由条件期望的唯一性即得所要之结论.

推论 7.5.4 设 $(\mathscr{X}^T,\mathscr{B}^T,P)$ 为概率空间,θ 为其上的平移算子,且保测 $\mathscr{G}_\theta$ 为 $\mathscr{B}^T$ 中的 θ-不变的子 σ-代数,即 $P(B\Delta\theta^{-1}B)=0$ $(\forall\ B\in\mathscr{G}_\theta)$. $A\in\mathscr{B}^T$. 令

$$I_n(u)=1_{\theta^n A}(u)\ (n\in T),$$

$$\nu_n(A)=\sum_{k=0}^{n-1}I_k,\ \pi(A)=E(I_0/\mathscr{G}_\theta),$$

则 $I=(I_n,n\in T)$ 为平稳 r. v. 列,且

$$\lim_{n\to\infty}\frac{1}{n}\nu_n(A)=\pi(A)\ (\mathrm{mod}P).$$

证 设 $m=m_1+m_2,m_i\geqslant 0\ (i=1,2)$,

$$f(u_0,\cdots,u_{m-1})=g_{m_1,m_2}(\theta^{-m_1}(u_0,\cdots,u_{m-1})),$$

则按推论 7.5.3,可得关于 g_{m_1,m_2} 的相应结论. 然后,构作一个非负函数列 $\{g_{m_1,m_2}:m_1\geqslant 0,m_2\geqslant 0\}$,使得

$$g_{m_1,m_2}\uparrow I_0(\mathrm{mod}P)$$

取极限后即得所要的结论.

定义 7.5.5 一般称推论 7.5.4 中的 $\pi(A)$ 为事件 A 的经验概率. 显然, $P(A)=E\pi(A)$. 若 $A\in\mathscr{G}_\theta$, 即 $P(A\Delta\theta^{-1}A)=0$, 则 $\pi(A)=I_0=1_A$. 因此, 若 $A=\varnothing$ 或 $\mathscr{X}^T(\mathrm{mod}P)$, 则 $\pi(A)$ 为常数(0 或 1).

定义 7.5.6 设 $(U,\mathscr{U},P)$ 为概率空间. S 为其上的保测变换. $\nu_n(A)=\nu_n(A,u)$ 表示序列 $\{u,Su,\cdots,S^{n-1}u\}$ 落到集 A 中的项数. 若对 $\forall\ A\in\mathscr{U}$, 有

$$\lim_{n\to\infty}\frac{1}{n}\nu_n(A,u)=P(A)\ (\mathrm{mod}P).$$

则保测变换 S 是遍历的. 若任意 S-不变集具有测度 0 或 1, 则称保测变换 S 是度量迁移的.

定理 7.5.3 概率空间 $(U,\mathscr{U},P)$ 上的保测变换 S 为遍历的充要条件是, 下列两条件之一成立:

(i) S 是度量迁移的;

(ii) $\lim\limits_{n\to\infty}\dfrac{1}{n}\sum\limits_{k=0}^{n-1}f(S^k u)=C\ (\mathrm{mod}P)$,

其中 C 为常数, f 是 P-可积函数.

证 证明路线: 遍历性 $\Rightarrow$ (i) $\Rightarrow$ (ii) $\Rightarrow$ 遍历性.

先设 S 遍历. 按定义 7.5.6, 有

$$\lim_{n\to\infty}\frac{1}{n}\nu_n(A,u)=P(A)\ (\mathrm{mod}P).$$

假如 $0<P(A)<1$, 其中 A 为某个 S-不变集. 则 $\{A,SA,S^2A,\cdots\}$ 中的元素仅区别一个 P-零集. 因此, $\nu_n(A)=n\cdot 1_A(\mathrm{mod}P)$, 即 $\dfrac{1}{n}\nu_n(A)=1_A(\mathrm{mod}P)$. 这与遍历性矛盾. 故条件(i)成立.

现设(i)成立, 即 S 为度量迁移变换. 设 f 可积. 令

$$f^*(u)=\lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1}f(S^k u)\ (\mathrm{mod}P),$$

则按定理 7.5.2, f^* 是 S-不变的. 从而集

$$A_x = \{u: f^*(u) < x\} \ (\forall\ x \in \mathscr{R})$$

是S-不变集. 按条件(i), $P(A_x)=0$或1. ($\forall\ x\in \mathrm{R}$). 这表明: $f^*(u)=C$(常数)(mod$P$). 故(ii)真.

最后设条件(ii)成立. 让 $f(u)=1_A(u)$. 按推论 7.5.1, $E(1_A/\mathscr{G}_S)=c$. 故

$$P(A) = EE(1_A/\mathscr{G}_S) = c.$$

$$\Rightarrow \lim_{n\to\infty} \frac{1}{n}\nu_n(A,u) = c = P(A) \ (\mathrm{mod}P).$$

按定义 7.5.6, S 是遍历的.

定义 7.5.7 设 $\xi=(\xi_n, n\in T)$是定义在$(\Omega,\mathscr{F},P)$上取值于$(\mathscr{X},\mathscr{B})$的 r. v. 列, P_ξ 是由 ξ 诱导的$(\mathscr{X}^T,\mathscr{B}^T)$上的概率测度. θ 为$(\mathscr{X}^T,\mathscr{B}^T,P_\xi)$上的平移保测算子. 若$\theta$是遍历的, 则称$\xi$为遍历 r. v. 列.

引理 7.5.5 设 θ 为$(\mathscr{X}^T,\mathscr{B}^T,P)$上的平移保测算子, θ 遍历的充要条件是: 下列两极限之一成立:

(i) $\lim\limits_{n\to\infty}\frac{1}{n}\sum\limits_{k=0}^{n-1}P(B\Delta\theta^{-k}A)=P(A)\cdot P(B)$;

(ii) $\lim\limits_{n\to\infty}\frac{1}{n}\sum\limits_{k=0}^{n-1}P(\theta^{-k}A/B)=P(A)$, 其中 $P(B)\neq 0$.

极限中的 $A,B\in\mathscr{B}^T$.

证 (i)和(ii)等价是因为

$$P(\theta^{-k}A/B) = \frac{P(B\cap\theta^{-k}A)}{P(B)}.$$

因此, 仅考虑(i)即可.

假如(i)成立. 则对$\forall$ θ-不变集 A, 有

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1}P(A\cap\theta^{-k}A) = P^2(A).$$

而 $P(A\cap\theta^{-k}A)=P(A)$. 故 $P(A)=P^2(A)$. 从而 $P(A)=0$ 或 1. 按定理 7.5.3, θ 是遍历的.

反之, 设 θ 是遍历的, 按推论 7.5.2, 有

$$\lim_{n\to\infty}\int_{\mathscr{X}^T}\frac{1}{n}\sum_{k=0}^{n-1}f(\theta^{-k}u)\cdot g(u)P(\mathrm{d}u)=\int_{\mathscr{X}^T}f^*(u)g(u)P(\mathrm{d}u).$$

令 $f(u)=1_A(u),g(u)=1_B(u)\ (u\in\mathscr{X}^T)$，则由 θ 遍历而有 $f^*(u)=P(A)\ (\mathrm{mod}P)$. 故

$$E(f^*\cdot g)=P(A)\cdot P(B).$$

即条件(i)成立.

定义 7.5.8 设 S 为$(\mathscr{X}^T,\mathscr{B}^T,P)$上的保测算子. 若对$\forall\ A$、$B\in\mathscr{B}^T$，且$P(B)\neq 0$，有

$$\lim_{n\to\infty}P(S^{-n}(A)/B)=P(A),$$

则称 S 为混合变换. 其条件称为混合条件.

引理 7.5.6 设 S 为$(\mathscr{X}^T,\mathscr{B}^T,P)$上的保测算子. S 为混合变换的充要条件是

$$\lim_{n\to\infty}E(f(S^{-n}u)\cdot g(u))=Ef(u)\cdot Eg(u),$$

其中 f、$g\in L^2$.

证 采用简单函数逼近的处理方法立即可得此结论.

定理 7.5.4 （强大数定律） 设 $\xi=(\xi_n,n\in T)$为定义在$(\Omega,\mathscr{F},P)$上的独立同分布的 r.v.列，且 $\xi_0\in L^1$. 则

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1}\xi_k=E\xi_0(\mathrm{mod}P).$$

证 显然，ξ 是平稳的. 按定理 7.5.2，有

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1}\xi_k=\xi^*(\mathrm{mod}P).$$

且 $E\xi^*=E\xi_0$. 同理，对$\forall\ m\geqslant 0$，有

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1}\xi_{k+m}=\xi^*(\mathrm{mod}P).$$

因此，ξ^* 是尾 σ-代数可测的. 按 Kolmogorov 0-1 律，有

$$\xi^*_1=c(\text{常数})\ (\mathrm{mod}P)$$

$$\Rightarrow\xi^*=E\xi_0(\mathrm{mod}P).$$

定义 7.5.9 设 $\xi=(\xi_n,n\in T)$为定义于$(\Omega,\mathscr{F},P)$上取值于

$(\mathscr{X},\mathscr{B})$的平稳序列. $\mathscr{F}_{\infty}=\bigcap\limits_{n}\mathscr{F}_{n}$,其中 $\mathscr{F}_{n}=\sigma(\xi_{n},\xi_{n+1},\cdots)$(称 $\mathscr{F}_{\infty}$为ξ的尾σ-代数). 若 $\mathscr{F}_{\infty}$中仅含概率为 0 或 1 的事件,则说 0-1 律适合于ξ.

定理 7.5.5 若 0-1 律适合于ξ,则定义在概率空间$(\mathscr{X}^{T},\mathscr{B}^{T},P_{\xi})$上的平移算子$\theta$是一个混合变换.

证 由于ξ平稳,θ为保测变换. θ的混合性等价于:对$\forall$ A、$B\in\mathscr{B}^{T}$,$P_{\xi}(B)\neq 0$,有

$$\lim_{n\to\infty}P_{\xi}(\theta^{-n}A/B)=P_{\xi}(A)$$

$$\Longleftrightarrow \lim_{n\to\infty}P_{\xi}(B\cap\theta^{-n}A)=P_{\xi}(A)P_{\xi}(B)\ (\forall\ A、B\in\mathscr{B}^{T}).$$

ξ的平稳性表明:$\mathscr{F}_{n}=\mathscr{F}_{-n}(\forall\ n\in T)$. 因此,$\mathscr{F}_{-\infty}$中的元素之概率为 0 或 1. 令$\eta_{n}=P(\tilde{B}/\mathscr{F}_{n})$,则易证

$$\eta_{n}\to P(\tilde{B}/\mathscr{F}_{-\infty})\ (\mathrm{mod}P)\ \text{当}\ n\to-\infty\ \text{时}$$

$$\Rightarrow\eta_{n}\to P(\tilde{B})\ (\mathrm{mod}P)\ \text{当}\ n\to-\infty\ \text{时}.$$

令$\mathscr{G}_{n}=\xi(\mathscr{F}_{n})$,$\tilde{B}=\xi^{-1}(B)$,则

$$P_{\xi}(B/\mathscr{G}_{n})\xrightarrow[n\to-\infty]{}P_{\xi}(B)\ (\mathrm{mod}P_{\xi}).$$

$$\Rightarrow P_{\xi}(B\cap\theta^{-n}A)=\int_{\theta^{-n}A}P_{\xi}(B/\mathscr{G}_{n})P_{\xi}(\mathrm{d}u)$$

$$\xrightarrow[n\to\infty]{}\lim_{n\to\infty}P_{\xi}(B)P_{\xi}(\theta^{-n}A)=P_{\xi}(B)P_{\xi}(A).$$

这里用到$\theta^{-n}A\in\mathscr{G}_{n}$.

定理 7.5.6 设$\xi=(\xi_{n},n\in T)$为定义于$(\Omega,\mathscr{F},P)$上的平稳 Gauss 序列,则在$(\mathrm{R}^{T},\mathscr{B}^{T},P_{\xi})$上的平移算子$\theta$当相关系数$R_{n}\to 0$时满足混合条件.

证 ξ的分布参数为:$m=E\xi_{0}$,(均值),$R_{n}=E(\xi_{n}-m)(\xi_{0}-m)$(相关系数),设

$$f(u)=f(u_{0},u_{1},\cdots,u_{p}),g(u)=g(u_{0},u_{1},\cdots,u_{p}).$$

这里f、g均为足够光滑的有界函数,且分别具有绝对可积的 Fourier 变换$f^{*}(\lambda_{0},\lambda_{1},\cdots,\lambda_{p})$,$g^{*}(\lambda_{0},\lambda_{1},\cdots,\lambda_{p})$. 则

$$Ef(\xi_{n},\cdots,\xi_{n+p})g(\xi_{0},\cdots,\xi_{p})$$

$$= E\int\cdots\int f^*(\lambda_0,\cdots,\lambda_p)g^*(\alpha_0,\cdots,\alpha_p)\exp\left\{i\sum_{k=0}^{p}(\lambda_k\xi_{n+k}\right.$$

$$\left.+\alpha_k\xi_k)\right\}\mathrm{d}\lambda_0\cdots\mathrm{d}\lambda_p\mathrm{d}\alpha_0\cdots\mathrm{d}\alpha_p$$

$$= \int\cdots\int f^*(\lambda_0,\cdots,\lambda_p)g^*(\alpha_0,\cdots,\alpha_p)$$

$$\cdot\exp\left\{-\frac{1}{2}\sum_{k,r=0}^{p}[(\lambda_k\lambda_r+\alpha_k\alpha_r)R_{k-r}\right.$$

$$\left.+\lambda_k\alpha_kR_{n+k-r}]\right\}\mathrm{d}\lambda_0\cdots\mathrm{d}\lambda_p\mathrm{d}\alpha_0\cdots\mathrm{d}\alpha_p.$$

因此，对$\forall\ p\geqslant 0$. 若$\lim\limits_{n\to\infty}R_n=0$，则

$$Ef(\xi_n,\cdots,\xi_{n+p})g(\xi_0,\cdots,\xi_p)\xrightarrow[n\to\infty]{}Ef(\xi_0,\cdots,\xi_p)Eg(\xi_0,\cdots,\xi_p).$$

于是，按引理 7.5.6，θ 是混合的.

注 2 混合条件含遍历性条件. 但反之不必真.

定理 7.5.7 设 ξ 为定义于$(\Omega,\mathscr{F},P)$上具有状态空间$(E,\mathscr{E})$的正常返不可约 Markov 链，具有不变初始分布 $v=(v_i,i\in E)$，这里 v 按定理 7.4.2 中定义，则由 v 和转移阵 $\mathscr{P}$ 决定的 Markov 逆链 ξ^* 是遍历的.

证 按§7.4 中的分析，逆链 ξ^* 是平稳的. 按引理 7.5.5，遍历性条件可表示为

$$\lim_{m\to\infty}\frac{1}{m}\sum_{n=0}^{m-1}P(\{\xi_k^*=i_k,0\leqslant k\leqslant s\}\cap\{\xi_{n+k}^*=j_k,0\leqslant k\leqslant r\})$$

$$=P(\{\xi_k^*=i_k,0\leqslant k\leqslant s\}\cap\{\xi_k^*=j_k,0\leqslant k\leqslant r\})$$

$$\Longleftrightarrow\lim_{m\to\infty}v_i\frac{1}{m}\sum_{n=0}^{m-1}p_{ij}^{(n)}=v_iv_j(\forall\ i,j\in E)$$

$$\Longleftrightarrow\lim_{m\to\infty}\frac{1}{m}\sum_{n=0}^{m-1}p_{ij}^{(n)}=v_j(\forall\ i,j\in E).$$

故 ξ^* 是遍历的.

参考文献

[1] Shiryayev, A. N. Probability, Springer-Verlag, 1984.

[2] Billingsley, P. Probability and measure, John Wiley and Sons. 1986.

[3] Chow, Y. S. and Teicher, H.
Probability Theory: Independence, Interchangeability, Martingales, Springer-Verlag. 1978.

[4] Billingsley, P. Convergence of Probability Measure. John Wiley and Sons. 1968.

[5] Gihman, I. I. and Skorohod, A. V.
The Theory of Stochastic Processes(I),
Springer-Verlag, 1979.

[6] Zheng Wei-an, A note on the Convergence of Sequences of Conditional Expectations of Random Variables. Z. W. v. G. 53, p291—292, 1980.

[7] Shiryayev, A. N. Optimal Stopping Rules,
Springer-Verlag, 1978.

[8] Snell, I. L. Applications of Martingales System Theorems,
Trans. Amer. Math. Soc. 73, p293—312, 1953.

[9] Liptzer, E. Sh. and Shiryayev, A. N.
Statistics of Random Processes(I), Springer-Verlag, 1977.

[10] Metivier, M. Semimartingales, de Gruytar, 1982.

[11] Chow, Y. S. and Robbins, H.
A martingale system theorem and applications,
Fourth Berkeley Symp. Math. Statist. Prob. 93—104, 1961.

[12] Chow, Y. S. and Siegmund, D.
Great Expectations: the Theory of Optimal Stopping,
Boston, 1971 (有中译本).

[13] Kemery, J. G., Snell, J. L. and Knapp.
Denumerable Markov Chains, Springer-Verlag, 1976.

[14] Meyer, P. A. Probability and Potentials. New York, 1966.

[15] Lehmann, E. L.

Testing statistical Hyperthesis, N. Y. Wiley, 1959. 1986

[16] Spitzer, F. Principles of Random Walk, Springer-Verlag, 1964.

[17] 中大编．测度与概率基础．广州:广东科技出版社,1983.

[18] 江泽坚、吴智泉合编．实变函数论．北京:人民教育出版社．1961.

[19] 严士健等著．概率论基础．北京:科学出版社．1982.

[20] 复旦大学编．概率论:第三册(随机过程)．北京:人民教育出版社,1981.

[21] F. 黎茨等著．庄万等译．泛函分析讲义第二卷．北京:科学出版社,1983.

[22] 胡必锦．论逆π-Bayes决策．数学物理学报,16(4):421—434. 1996.

[General Information]
书名=半鞅序列理论及应用
作者=胡必锦著
页数=373
SS号=11105146
DX号=
出版日期=1997年03月第1版
出版社=华中理工大学出版社